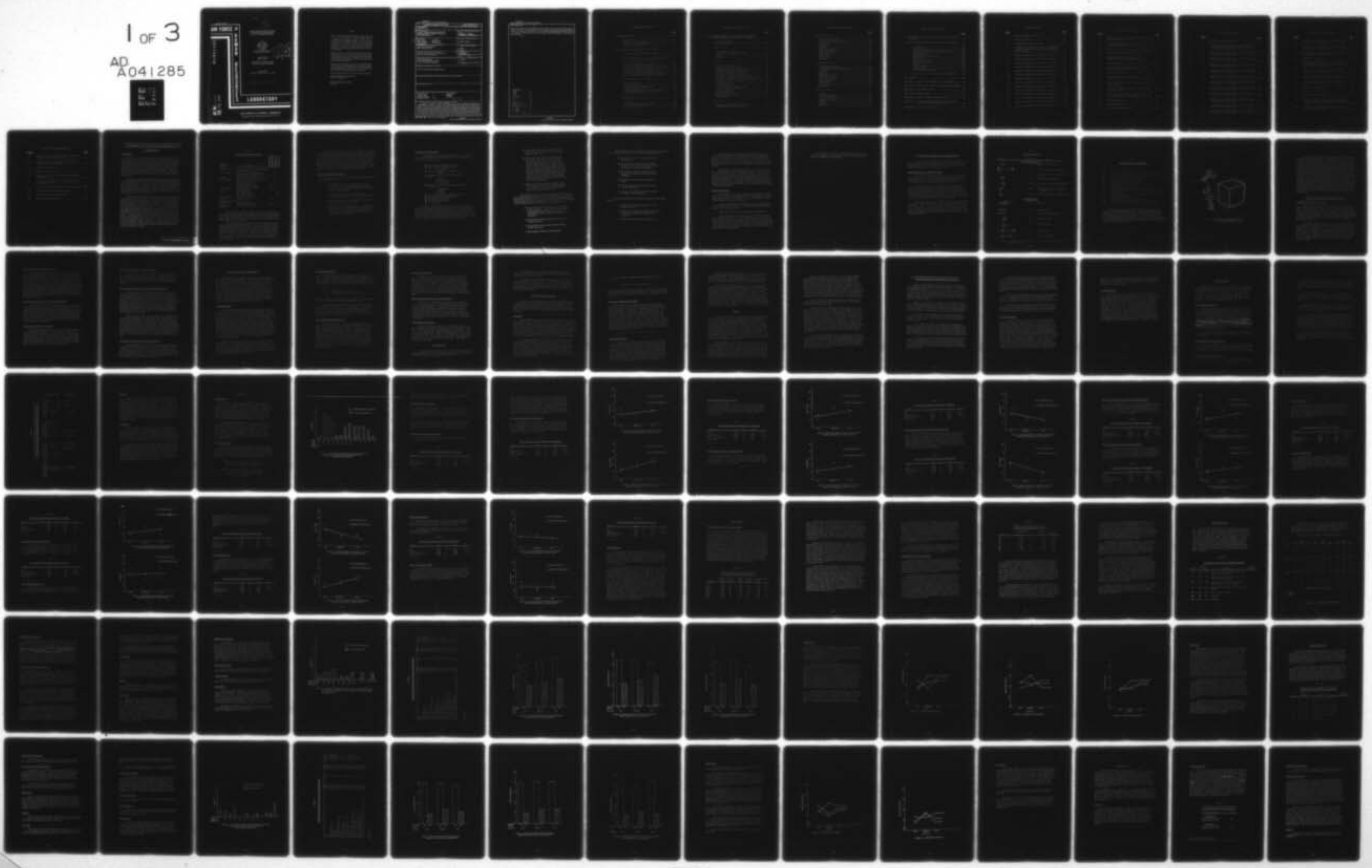
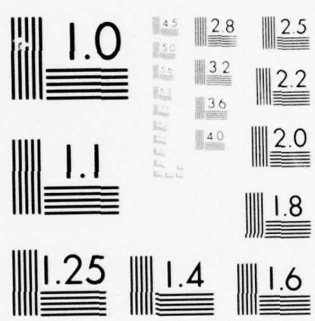
AD-A041 285 APPLIED PSYCHOLOGICAL SERVICES INC WAYNE PA F/G 5/7
STUDIES AND DESIGN SPECIFICATIONS FOR COMPUTERIZED MEASUREMENT --ETC(U)
OCT 76 A I SIEGEL, A R WILLIAMS, W J LAPINSKY F41609-75-C-0037

UNCLASSIFIED AFHRL-TR-76-77 NL

1 OF 3

AD A041285





MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

AFHRL-TR-76-77

12

B.S.

AIR FORCE HUMAN RESOURCES LABORATORY

AD A041285

AD No.
DDC FILE COPY

STUDIES AND DESIGN SPECIFICATIONS FOR COMPUTERIZED MEASUREMENT OF TEXTUAL COMPREHENSIBILITY

By

Arthur I. Siegel
Allan R. Williams
Walter J. Lapinsky
Tom A. Warms
J. Jay Wolf
Applied Psychological Services, Inc.
Wayne, Pennsylvania 19087

Steven D. Groff
James R. Burkett

TECHNICAL TRAINING DIVISION
Lowry Air Force Base, Colorado 80230

October 1976
Final Report for Period March 1975 – June 1976

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

This final report was submitted by Applied Psychological Services, Incorporated, Wayne, Pennsylvania 19087, under contract F41609-75-C-0037, project 1121, with Technical Training Division, Air Force Human Resources Laboratory (AFSC), Lowry Air Force Base, Colorado 80230. Captain Steven D. Groff, Instructional Technology Branch, was the contract monitor.

This technical report has been reviewed and is approved for publication.

MARTY R. ROCKWAY, Technical Director
Technical Training Division

DAN D. FULGHAM, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFHRL-TR-76-77	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) STUDIES AND DESIGN SPECIFICATIONS FOR COMPUTERIZED MEASUREMENT OF TEXTUAL COMPREHENSIBILITY		5. TYPE OF REPORT & PERIOD COVERED Final March 1975 – June 1976 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Arthur I. Siegel, Allan R. Williams, Walter J. Lapinsky, Tom A. Warms, J. Jay Wolf, Steven D. Groff, James R. Burkett		8. CONTRACT OR GRANT NUMBER(s) F41609-75-C-0037
9. PERFORMING ORGANIZATION NAME AND ADDRESS Applied Psychological Services, Inc. Wayne, Pennsylvania 19087		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62703F 11210412
11. CONTROLLING OFFICE NAME AND ADDRESS HQ Air Force Human Resources Laboratory (AFSC) Brooks Air Force Base, Texas 78235		12. REPORT DATE October 1976 13. NUMBER OF PAGES 254
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Technical Training Division Air Force Human Resources Laboratory Lowry Air Force Base, Colorado 80230		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

cognitive processes
comprehensibility
information processing
intellective processes
psycholinguistics
readability
text analysis

A previous report (AD-A 001 537)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A previous report (Siegel & Burkett, 1974) defined a series of 14 novel measures for determining the comprehensibility of English text on the basis of current psycholinguistic and Structure-of-Intellect oriented concepts. That report not only suggested the potential usefulness of the measures, but also conjectured the feasibility of automating the calculation of these measures. The ~~current~~ report ~~builds upon~~ the prior work and takes the next logical steps in implementing these measures for computer application. First, these measures are analytically defined and described. Then, selected measures are subjected to "laboratory" experimental investigation using Air Force Manuals, Career Development Course materials, and USAF Technical Orders as sample texts. The results of these experiments are presented. An automatic calculation method is then developed for each of the 13 selected

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Item 20 (Continued): fr p1473A)

measures. The structure of the programming specifications is modular and is intended to calculate the measures for variable size blocks of texts. Flow charts and summary descriptions of the program attributes are also presented, together with explanations of run request syntax, sample measures calculations, and output formats. This report then constitutes a complete definition of the program suitable for future implementation on an automatic data processing system.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont.)

TABLE OF CONTENTS (cont.)

TABLE OF CONTENTS (cont.)

LIST OF TABLES

LIST OF TABLES (cont.)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (cont.)

LIST OF ILLUSTRATIONS (cont.)

STUDIES AND DESIGN SPECIFICATIONS FOR CONCEPTUALIZED MEASUREMENT OF TEXTUAL COMPREHENSIBILITY

I. INTRODUCTION

Background

For several years, Applied Psychological Services, under the sponsorship of the Air Force Human Resources Laboratory, has conducted research into methods for measuring and increasing the comprehensibility of written text (particularly textual materials utilized during the course of Air Force technical training). The long range goal of this work is to identify and develop a method which will facilitate the comprehension of textual materials employed in the training situation. Achievement of this capability could be expected to reduce training time and costs, and to increase training effectiveness.

The efforts toward this long term goal initially concentrated on identification of methods previously employed in measuring comprehensibility (Williams, Siegel, & Burkett, 1974) and on the acquisition of data relative to the questions of how, and in what training context, auditory supplementation of written materials would increase the transfer of knowledge (Siegel, Lautman, & Burkett, 1974). A major conclusion of these efforts was that the new methods for measuring comprehensibility are required.

Williams et al. (1974) noted that a long list of techniques for calculation of comprehensibility measures had been offered for consideration over the past 30 years. A sample of those considered to be of principal interest is contained in Table 1. For convenience, Table 1 groups the measures into three classes: structural complexity, word divergency, and parts of speech. These deal principally with what one might call mechanically oriented factors. They deal with quantities of words, sentences, syllables and their rates of occurrence, but are not concerned with meanings of words or phrases per se. They have been in use for some time, not only because they measure reading difficulty in some sense, but also because they are suitable for relatively easy calculation by hand. These measures have been used principally to determine the reading grade level (RGL) of text. However, a major shortcoming of these measures is that they fail to consider the inherent mental or intellectual load placed upon the readers by the text.

Table 1

Selected Readability Measures

Author or Developer	Variable	Structural Complexity	Word Divergency	Parts of Speech
Lively & Pressey	Median index no. of sampled words on Thorndike word list		x	
Dale & Tyler	No. different technical terms		x	
	No. different non-technical terms		x	
	No. independent clauses	x		
Thorndike	Proportion of words not on word list		x	
Gray & Leary	No. of words not on Dale list		x	
	No. personal pronouns			x
	Sentence length	x		
	Percent different words		x	
	No. prepositional phrases	x		
Dale & Chall	Words not on Dale list		x	
	Sentence length	x		
Gillie	No. finite verbs			x
	No. definite articles			x
	No. abstract nouns			x
Powers, Sumner	Word length in syllables	x		
& Kearl	Sentence length	x		
[Revision of Flesch]				
Smith & Senter	Sentence length	x		
[ARI]	word length (letters)	x		
Caylor et al. [FORCAST]	No. one syllable words	x		

As pointed out by Bormuth (1966), until very recent years no theoretical base was available from which to generate testable hypotheses relating to readability. Powerful theories of language behavior did not exist, so that only the most obvious statistical characteristics of the written text were studied.

This void has begun to be filled by modern linguistic and psycholinguistic research and by models of overall intellective functioning such as the Structure-of-Intellect of Guilford and his associates (Guilford, 1967; Guilford, Comrey, Green, & Christensen, 1950; Guilford, Geiger, & Christensen, 1954; Guilford & Hoepfner, 1966; Hoepfner, Guilford, & Merrifield, 1964). The readability factors examined in the current series of studies are based on implications of the Structure-of-Intellect model and of experimental psycholinguistic findings.

In earlier work in the present series, the effects of psycholinguistically oriented and intellective function related variables on the comprehensibility of text were demonstrated, and methods for improving readability were defined (Siegel & Burkett, 1974). Fourteen specific comprehensibility measures were specified, and the feasibility of developing a computer technique to automate their calculation was determined to have only a modest associated risk. Siegel and Burkett concluded that: "The results supported a contention favoring the potential of psycholinguistic and intellective concepts for readability/comprehensibility measurement."

Purpose of the Present Program

The current research and development effort had as its focus the following objectives:

1. formulation of specific procedures for measuring each of a set of psycholinguistically and Structure-of-Intellect oriented variables for measuring textual comprehensibility

2. selection of measures most appropriate for use in automatic computation of a comprehensibility index

3. experimental study of the power and characteristics of the comprehensibility variables

4. development of specifications for a computer program, called the Comprehensibility Measures (CM) program. When implemented, CM would process blocks of texts and display the results of the comprehensibility measurements.

Utilization of the Technique

Eventual application of the automatic analytic technique was held in mind throughout the present program. A summary list of potential uses is given below:

- determine comprehensibility of text
- isolate causes of low/high scores
  - by measure
- determine placement of a written work
  - by RGL
- compare comprehensibility of two or more works
  - by author
  - by time period
  - by type of work
- compare comprehensibility between portions of a work
  - by paragraph
  - by page
  - by chapter
- writer training (author diagnostic assistance)
- inter-author discrimination studies
- statistical data generation
- sentence parsing studies
- screening text

Consider a writer who is concerned with the comprehensibility of his material and who has completed a written work in draft form. He would like to submit the material to a comprehensibility measurement analysis. Assume he has access to the computer system for which the CM program will have been prepared. This access could be either through the submittal of computer run requests to the computer center or via a remote terminal. He would then:

- arrange for the text to be prepared in machine readable form (e.g., punched cards, magnetic tape, etc.)

- select values for a variety of parameters and options which describe the run (these are discussed in Chapter V and in more detail in Appendices A and B of the present report. They include, for example, the size of the text blocks into which the total text is to be subdivided for measurement. The CM program provides for determination of measures for each text block whose size may be specified in terms either of a prespecified number of words or identification of the start of prespecified blocks)

- request a computer run either: (1) to check the text against a currently available dictionary, or (2) to calculate and display comprehensibility measures themselves.

The dictionary check would usually be performed first because the calculation of certain measures depends on a prespecified percentage of words in the text actually appearing in the dictionary. If the dictionary check is requested, the requester will be presented with the following types of information as output from CM:

- a list of words in the text which do not appear in the dictionary. These would indicate either a misspelling which should be corrected, or that additional entries should be made to the dictionary

- the location of the first occurrence of such words in the text

- the number of occurrences of each of these words in the text

- percentage of words not in dictionary.

If the request is for a computation of measures, the following will be made available at several levels of detail as desired:

- the value of each of 14 measures for each block of text
- the measured reading grade level (RGL) using up to three of the "classical" (prior) methods, for comparison purposes
- the value of a variety of detailed counts used in the computation of the measures and RGLs
- a block by block summary of measure values
- a single composite value of the 14 measures for each block
- averages over all blocks of each measure, the RGLs, and the composite.

These data present the text writer with indices with which he can:

1. compare the comprehensibility of current and prior works
2. compare the comprehensibility of individual blocks (or paragraphs or chapters) with each other
3. isolate the "cause" of low overall scores.

This will then specify blocks for revision, if desired. The process may be repeated or continued for succeeding chapters or sections of the text as they are prepared, or for revised portions as they are rewritten or edited. In this way, the writer has an iterative and, in a sense, a corrective procedure at his disposal so that he may continue to strive for comprehensibility scores which are acceptable.

An alternate use of the CM program is to assist an evaluator in assessing the comprehensibility of a report, manual, lesson, or other written work; i.e., essentially a quality control function. Consider the case in which many such documents are to be evaluated, and it is not feasible to encode completely the entire text into machine readable form. Then, the evaluator could sample several blocks from the text to be evaluated. The CM program would then evaluate each block and show block as well as average per block data over the entire sample.

Scope of the Report

Siegel and Burkett (1974) defined and investigated 14 comprehensibility measures. In Chapter II of the present report, some of these measures have been redefined so as to facilitate their implementation within a computer program.

Chapter III describes the methods and results of additional experiments performed to gain insight into the operating characteristics of the various measures.

Chapter IV presents conclusions and recommendations.

The various aspects and requirements for the CM program are contained in Chapter V. A special dictionary is cited as a requirement for input to this program, and its features are described. A description of each module of the program is also presented, along with summary logic flow diagrams and run request information. Output results and their formats are specified in sufficient detail to allow initiation of programming.

The Appendices contain complete guidelines for the development of a digital computer program which will allow automatic analysis of textual comprehensibility.

II. SELECTION AND DEFINITION OF MEASURES

The Structure-of-Intellect and the psycholinguistically oriented comprehensibility measures are presented in Tables 2 and 3, which show for each of 14 measures the name, abbreviation, and associated formula used to calculate its value.

Guilford's Structure-of-Intellect Model

The Structure-of-Intellect (SI) model developed by Guilford and his associates (Guilford, 1967; Guilford et al., 1950, 1954, 1966; Hoepfner et al., 1964) produced a hypothetical construct as to the nature and structure of human intellective activity.

The SI model is a cross-classification representation that classifies intellectual abilities along three different dimensions. Each dimension is divided into categories which intersect with the categories of the other dimensions of ability. Mental operations represent one dimension of classification in the SI model. The five mental operations are: (a) cognition, (b) memory, (c) divergent production, (d) convergent production, and (e) evaluation.

The second classification dimension of the SI model involves the content areas of information on which the mental operations are performed. These areas of information include: (a) figural, (b) symbolic, (c) semantic, and (d) behavioral. Twenty separate abilities can be derived from the combination (intersection) of the five categories in the mental operation dimension and the four categories in the contents dimension.

Table 2

Summary Measures

Formula	*Abbreviation*	*Name*
Structure-of-Intellect		
$1 - \frac{NDWB}{TNWB}$	CMU	Cognition of Semantic Units
$\left(\frac{NSNB}{TNSB - 1} + NORB\right) / TNWB$	CMR	Cognition of Semantic Relations
$\frac{TNAB}{4}$	NMS	Convergent Production of Semantic Systems
$1 - \frac{NDNB}{TNWB}$	MMU	Memory of Semantic Units
$1 - \frac{NSWB}{TNWB}$	ESI	Evaluation of Symbolic Implications
$\frac{TNWB}{TPSB}$	NMI	Convergent Production of Semantic Implications
$\frac{TNEB}{TPSB}$	DMU	Divergent Production of Semantic Units
Psycholinguistic		
$\frac{TNWB}{TNSB \cdot \sum_{W} YDS}$	YD	Yngve depth
$TNWB/TNMB$	MD	Morpheme depth
$\sum_{S} TCS/TNCB$	TC	Transformational complexity
$1 - \frac{\sum NNPS}{TNSB}$	CE	Center embedding
$1 - \frac{\sum_{S} NCLS}{TNSB}$	LB	Left branching
$TNSB/\left[\sum_{S} NCRS + TNSB\right]$	RB	Right branching
$1 - \left[\sum_{S} DCS/TNSB\right]$	DC	Deleted complements

NOTE: If a measured value exceeds the 0 to 1 range, it is bounded at the appropriate limit.

Table 3

Symbology for Table 2 Formulae

NDWB	Number of different words in a text block
TNWB	Total number of words in a text block
NSNB	Number of shared nouns, count of nouns in adjacent sentences
NORB	Number of references (number of pronouns) in a text block
NDNB	Number of different nouns in a text block
NSWB	Number of abbreviated or symbolic words in a text block
NPPB	Number of potential parses per sentence
TNMB	Total number of morphemes per block
TNEB	Total number of elucidations per block
TNSB	Total number of sentences per block
TNWS	Total number of words per sentence S= 1, ..., TNSB
YDS	Yngve depth of a sentence
TCS	Transformational complexity value for a sentence
NNPS	Number of noun phrases to the right of the subject verb in a sentence
NCRS	Number of modifying clauses on the right of the object noun phrase of a sentence
DCS	Deleted complement in a sentence
TNCB	Total number of clauses in a text block
TPSB	Total number of parts of speech in all words of a block
NCLS	Number of chained modifying clauses on the left of subject noun
TNAB	Total number of aids to comprehension in a text block

The final dimension of intellect in the SI model concerns the formal types of information dealt with. These informational types or products can be units, classes, relations, systems, transformations, and implications. When the six products are combined with the five operations and with the four contents, 120 cells result. The total model is composed of these 120 abilities and can be viewed as a three-dimensional cube, shown in Figure 1.

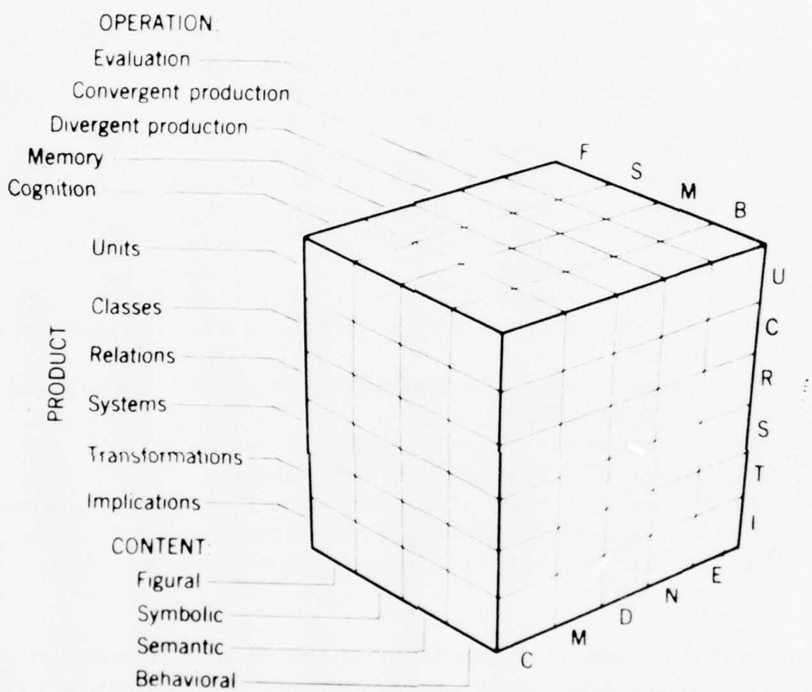


Figure 1. The Structure-of-Intellect model (from Guilford & Hoepfner, 1971).

Our hypothesis is that textual materials which require high levels of SI oriented abilities for mastery can be said to be less comprehensible than materials which require lower levels of these abilities. The defined metrics can be applied to textual materials to reflect the SI oriented abilities required to master the materials. This involves adapting the model to a comprehensibility format such that the degree to which a particular reading selection is loaded in various factors may be quantified. Since the model contains 120 cells (abilities), a sample was required. To this end, those eight abilities which seemed most relevant to the comprehensibility problem in the Air Force technical training context were selected. Those selected were: cognition of semantic units, cognition of semantic relations, memory of semantic units, memory of figural units, convergent production of semantic implications, convergent production of semantic systems, divergent production of semantic units, and evaluation of symbolic implications. Each of these is discussed below; however, for precise definitions of the variables used in the experiments described in Chapter III, the reader should refer to Tables 2 and 3.

Factors Derived from the SI Model Involved in Comprehensibility

Cognition of Semantic Units (CMU)

CMU in the comprehensibility context involves the extent to which the text forces the reader to recognize a diversity of word forms. Thus, the rhyme, "One little piggy went to market, one little piggy stayed home" is held to be readable because of the common word usage. The redundancy of words is held to increase readability. The same material written as "A unitary small piggy went to market, one little hog stayed home" is considered to be less readable than the original text.

According to Guilford (1967), cognition of semantic units is most directly measured by vocabulary tests. In the context of comprehensibility, an analogous measure is the ratio of the number of different words (types) to the total number of words (tokens). The CMU measure is equal to the inverse of the type-token ratio, subtracted from one. Measures of vocabulary diversity are commonly found in existing readability formulas; the effect of vocabulary diversity on comprehensibility was also demonstrated by Siegel and Bergman (1974).

Cognition of Semantic Relations (CMR)

CMR was defined as the extent to which the text forces the reader to recognize the relationship between two items or words. Guilford (1967) used analogy and word linkage tests to measure this factor. In word linkage tests, the test taker is required to match sets of words in terms of their relatedness or connectedness. This factor is believed to be reflected in reading material when the reader must form analogies or word linkages while reading. One would expect that increasing the requirement for relational thinking in a reading selection would decrease reading comprehension. Thus, if assuming that the demand for relational thinking is a function of the number of common words and references between adjacent sentences, the fewer shared common words and references between sentences, the more difficult the text.

Convergent Production of Semantic Systems (NMS)

Guilford (1967) describes convergent production of semantic systems as the ability to order or organize material. Siegel and Bergman (1974) were able to manipulate comprehension by varying organization of the number of comprehension aids which appear per standard block of text. Examples of comprehension aids are mnemonic devices or a figure. NMS was considered for investigation although not included in the CM program. NMS was calculated as the number of aids to comprehension in a block divided by four--an arbitrarily selected constant.

Memory for Semantic Units (MMU)

According to Guilford (1967), memory for ideas is most directly addressed by memory for semantic units. Siegel and Bergman (1974) demonstrated that replication of facts increased comprehensibility. The analogous measure involves the number of different nouns appearing in a passage, divided by the number of words in the passage. It was assumed that repeated or related facts will be presented using the same noun as subject, allowing that measure to serve as an index of fact repetition.

Evaluation of Symbolic Implications (ESI)

Guilford (1967) used abbreviations tests to measure evaluation of symbolic units. A corresponding comprehensibility measure involves the frequency of occurrence of abbreviations in passages of running text. Siegel and Bergman (1974) demonstrated that frequency of occurrence of abbreviations influenced comprehension.

Convergent Production of Semantic Implications (NMI)

Guilford and Hoepfner (1971) used syllogisms, attribute listing, missing links, and sequential association tests to measure convergent production of semantic implications. Reading material loaded in convergent production theoretically requires the reader to perform syllogistic reasoning tasks. Material which does not require this ability completes the syllogism for the reader. By increasing the demand for convergent production of semantic implications a text should become less comprehensible. Siegel and Bergman (1974) found that textual material which imposed syllogistic reasoning on the reader was less comprehensible than that not involving such reasoning.

Identification of syllogisms by computer was considered beyond current capabilities. Accordingly, in the current work, a measure was employed in which the number of words in a passage was divided by the total of the dictionary-listed numbers of parts of speech of the words of a textual sample. This was considered to be an index of the number of possible parses of a sentence, and was thus analogous to a measure of syllogistic reasoning. That it, it is assumed that a sentence becomes more ambiguous as its number of possible parses increases. The increased load on the reader is assumed to increase the syllogistic load required to decode the sentence.

Divergent Production of Semantic Units (DMU)

According to Guilford (1967), DMU involves the ability to enumerate class members given certain class properties. With regard to the comprehensibility of training texts, DMU relates to the demand upon a reader to enumerate class members on his own rather than have the class member supplied in the reading selection. Siegel and Bergman (1974) demonstrated a relationship between the presentation of example(s) and comprehension.

Psycholinguistic Measures

The second set of measures was based on concepts drawn from the psycholinguistic literature. The literature and prior studies in this series (Siegel & Burkett, 1974) have suggested a number of rules for making a sentence more comprehensible. Some of these suggestions were to: (1) decrease word depth (Bormuth, 1969; Foss & Crains, 1970), (2) decrease morpheme depth (Bormuth, 1969), (3) change passive to active voice when there is a possibility of a reversal of subject and object; this reduces structural as well as semantical problems (Gough, 1965; Slobin, 1966; Fodor, 1971), (4) avoid center embedding whenever possible (Schwartz, Sparkman, & Deese, 1970; Wang, 1970), (5) avoid right-branching sentences whenever possible (Schwartz et al., 1970), and (6) write affirmative sentences when possible (Gough, 1965; Slobin, 1966).

Yngve Depth (YD)

Yngve (1960) developed a model of sentence production which claimed that a person produces sentences by generating a "sentence structure tree" in a top to bottom-left to right direction. Accordingly, at any given time, a speaker has produced only that portion of the left-hand side of the tree necessary to produce the word spoken. As the speaker works down the tree, he produces both branches of a node, but must store the right branch in memory while expanding the left branch. Bormuth (1969) found that sentence depth was correlated with the difficulty of a passage. Martin and Roberts (1966) held sentence length constant and varied the Yngve depth of sentences and found that sentences of lesser complexity were recalled significantly more frequently than sentences of greater structural complexity. The finding that mean linguistic depth is a strong predictor of sentence comprehensibility was replicated by Wang (1970) and by Lambert and Siegel (1974).

YD was computed in the present work by dividing the total number of words in a textual block by the product of the number of sentences in the block and the sum of the sentences' YD values. The YD value for a sentence was determined by parsing and determining the number of right branches in the sentence which had to be anticipated by a reader or listener. This value, then, is a measure of the overall depth of the sentences in the block.

Morpheme Depth (MD)

A morpheme is the meaning carrying unit of language. It does not always correspond to the syllable. For example, the word "flower" is one morpheme. Bormuth (1969) speculated that the comprehensibility of an individual word could be dependent on how many morphemes are "buried" within it. For example, in the word

un/happi/ness:

```
un    = morpheme denoting "not"
happi = morpheme denoting a state of mood
ness  = morpheme denoting a condition or quality
```

A person reading this word must have knowledge of the meaning of all three morphemes in order to comprehend the word.

Lambert and Siegel (1974) found the mean number of morphemes per word to be related to comprehensibility. Accordingly, MD was defined as the number of words sampled divided by the number of morphemes in the sampled words. As for all of the measures, low values of this index were expected to be associated with passages of low comprehensibility.

Transformational Complexity (TC)

According to theories of transformational grammar, sentences of any type or level of complexity are produced, or interpreted, through transformations relative to simple, active "kernels." Interpretation of passive, negative, or passive negative sentences or independent clauses requires successively more, or more elaborate, transformations from the basic active kernel. Lambert and Siegel (1974) found the relative frequency of errors of interpretation of sentences of these four classes to follow the order described above. The present measure of TC was based on assignment of a score based on the error rate of interpretation of sentences of the corresponding type observed by Lambert and Siegel. The average of these scores per sample of text provides the TC index.

Center Embedding (CE)

Schwartz et al (1970) demonstrated that the inclusion of phrases or clauses between a sentence's subject and its predicate decreased comprehensibility. This finding was confirmed by Lambert and Siegel (1974). To measure the degree of center embedding in a block of text, a measure was devised in which the number of phrases (prepositional, adverbial, etc.) and clauses (relative, adverbial, verb, complement, etc.) between the subject(s) and predicate(s) of a sample of sentences is divided by the number of sentences examined to yield a measure of comprehensibility.

Right Branching (RB) and Left Branching (LB)

Schwartz et al. (1970) also found that addition of phrases or clauses to the left end of a sentence (prior to the subject) decreased comprehensibility. Addition of similar material following the sentence predicate did not degrade comprehensibility. Lambert and Siegel (1974) tested these variables and obtained mixed results. These sentence structural aspects are evaluated in the present study through measures which consider the frequency of right or left branching, respectively, within a sampled group of sentences.

Complement Deletion (DC)

Certain surface structure features of sentences serve to mark the deep, or meaning carrying, structure of sentences. The word "that" when used as a complement as in "He said that I should go," serves that function. Hakes (1972) found the inclusion or deletion of this complement to affect comprehensibility. The measure, in the present study, which reflects this variable is the ratio of the number of occurrences of deleted complements to the number of sentences in a passage.

Integration

These psycholinguistically and SI oriented measures were selected for inclusion in the CM program on the basis of the conjecture that textual comprehensibility is a function of each.

All formulas have been scaled so that high values indicate improved comprehensibility and lower values imply a poorer comprehensibility on a scale of 0 to 1.

A companion report (Williams, Siegel, Burkett, & Groff, in press) presents the results of work in which norms for each measure and for a variety of types of documents were determined. For completeness, the resultant tables are included here as Appendix G. These data were used in the CM program to enable the common scaling of all calculated measures

Reading Grade Levels

As part of the CM program, two classical reading grade level (RGL) indices and one readability index are calculated. The RGLs estimate the school grade level of the given text. These are to be products of the CM program since the basic counts for these indices will be available from the CM calculation sequences or will be readily obtained.

FORCAST

The FORCAST formula was developed to measure the readability of Army technical literature (Caylor, Sticht, Fox, & Ford, 1972). The authors considered existing formulas inappropriate for their purpose because school students and school or general texts had been employed in developing prior readability formulas. This type of standardization was believed to make suspect the applicability of prior formulas to technical publications for adults. Moreover, application of many of the existing formulas requires special grammatical or linguistic competence on the part of the person attempting to apply the formulas.

A literature search provided Caylor et al. (1972) with a list of 15 structural properties of text that had been applied in previous readability formulas and required no special competence or equipment to measure. Correlations between cloze score and each of the structural properties were computed and regression equations were derived. Their preferred formula employed only a single factor, number of one syllable words per passage. This factor is very easily measured, and the addition of other factors yielded no practical increase in predictive power.

The FORCAST formula predicts reading grade level as equal to:

$$20 - (\text{number of one syllable words}/10)$$

The correlation between predicted RGL of a passage with tested RGL associated with 35 percent cloze score was .87. A subsequent application using new test passages and new subjects produced a correlation of .77.

Automated Readability Index (ARI)

Smith and Senter (1966) developed a readability equation whose data may be collected from mechanical counters easily installed on an electric typewriter. This technique allows measurement of readability at essentially rough draft typing speed. Mechanical counters are used to record the number of key strokes, blank spaces, and sentences (an equals sign must be typed at the end of each sentence; the number of activations of this key indicates the number of sentences typed). From these counts, the mean number of words per sentence [number of spaces divided by number of sentences (w/s)] and the mean length of words [number of strokes divided by number of spaces (s/w)] may be computed. Based on examination of grade school texts, the regression equation predicting grade level from the above ratios is:

$$\text{ARI RGL} = 0.50(w/s) + 4.71(s/w) - 21.43.$$

Flesch Reading Ease

Probably the most popular readability formula developed to date is that developed by Rudolf Flesch. Working in the 1940's, Flesch concluded that sentence length is important to predicting comprehension for adult readers. He similarly indicated that the readers' interest in a topic should also be related to readability. His "reading ease" formula predicts arbitrarily scaled "reading ease" as a function of word length of sentences and number of syllables per 100 words. His "human interest" index, also scaled arbitrarily, is based on rates of occurrence of personal words and of sentences addressed to the reader (Flesch, 1948).

The Flesch formulas, especially the "reading ease" formula, have become the most widely applied in the entire history of readability research. This wide application is due in part to the ease of computation of his formulas and partly to the wide exposure given to his formulas through a long series of popularized books.

Flesch based his "reading ease" formula on data based on a large set of reading passages normed in terms of comprehension test score and corresponding school grade level. This set of original passages, developed in 1926, was revised in 1950 to reflect modern topics and changes in population reading levels. The "reading ease" score was recomputed by Powers, Sumner, and Kearl (1958), based on the newly available data. According to their revision, arbitrarily scaled "reading ease" is equal to -2.2029 + (.0778) (mean sentence length) + (.0455)(number of syllables per 100 words). This recent form of the "reading ease" formula has been implemented in the CM program specification.

Cloze

In the experimental work reported in subsequent sections of the present report, cloze score was employed as the readability criterion. The cloze procedure, introduced by Taylor (1953), was demonstrated to rank standard reading passages in the same order as did other readability formulas. In the cloze procedure, readers are presented with samples of text, from which some words are deleted and replaced by blank spaces. The readers are requested to fill in the blank spaces with the words they think were deleted. To the extent that the author uses the words that the reader expects and understands, the reader will fill in the correct words. The technique assumes that readability is a direct function of the number of omitted words which the reader is able to fill in.

Taylor indicated that the ordering of cloze scores is maintained regardless of the system employed in word deletion. He used four different deletion systems on three standard passages, each of a different difficulty level, and found that the rankings were the same despite the deletion system employed. The four deletion systems involved deleting every fifth word, every seventh word, every tenth word, and 10 percent of the words at random. Others have reported that deleting 20 percent of a passage will yield sensitive measures.

The cloze method is free from many of the disadvantages of the traditional readability measures. It can be applied more appropriately to highly technical and unusual materials. Very technical material might be rated as difficult by the Flesch formula, but not by the cloze technique, if the subjects reading the passage were trained in the subject matter area. Conversely, the readability of the writings of authors such as Gertrude Stein, who write in short sentences with relatively simple vocabulary, but whose style is such that the material is difficult might not be accurately reflected by the Flesch and Dale-Chall techniques. The cloze test might be expected to reflect accurately the difficulty of such reading passages.

Taylor reported correlations of .70 and .80 between cloze scores and comprehension scores received by Air Force trainees on Air Force technical material.

In developing their readability equation, Caylor et al. used cloze score as the criterion of readability. Because they believed the cloze test to be more objective than multiple choice tests or the other more traditional indices of comprehension. They also pointed out that cloze had "consistently yielded very high correlations with multiple choice tests and other more subjectively constructed measures of comprehension and difficulty" [Caylor et al., 1972, p. 12]. Additionally, as part of their work, they found a correlation of approximately .80 between cloze score on 150-word passages chosen from the readings required in a wide range of Army jobs and achieving reading grade level, as measured by the United States Armed Forces Institute Reading Achievement Test III, Form A, Abbreviated Edition.

The inherent advantages of the cloze procedure are: (1) scoring ease, (2) scoring reliability, (3) ease of application to nonstandard material, and (4) accounting for the reader's interest in and prior knowledge of the content. The disadvantages of the procedure are: (1) cloze is a measure of readability, not a predictor of readability, (2) a sizeable sample of subjects is required, and (3) it may not reflect all types of comprehension.

III. FURTHER VERIFICATION AND ELABORATION OF COMPREHENSIBILITY ASSESSMENT VARIABLES

The SI oriented variables included in the computer algorithm were originally developed and evaluated for influence on readability by Siegel and Bergman (1974). The psycholinguistic variables were similarly developed and examined by Lambert and Siegel (1974). In the cited studies, all SI based variables were found to be related to comprehension. The relationships between the psycholinguistic variables and comprehension were less clear cut, but the measurement concepts remained viable.

However, it seemed that further test and evaluation of both the SI and the psycholinguistically oriented variables were warranted. Such work would allow additional evaluation of the variables included in the CM program, as well as the development of insights relative to interaction effects with which the prior work was not concerned.

In the study of psycholinguistic variables performed by Lambert and Siegel (1974), variables were presented in individual sentences in many cases, and a variety of comprehension measures were used.

In the related exploratory SI oriented work of Siegel and Bergman, completion, true-false, and short answer questions were used to evaluate comprehension. The variables were presented in paragraphs prepared for the purpose, but Air Force technical training materials were not involved, and the range of the variables was not the same as that found in Air Force technical training materials.

In the first three of the four studies described in the present chapter, all passages were developed through modification of current Air Force materials; all variables approximated the range of the variables found in Air Force technical training and related materials; and a single criterion measure of comprehension, the cloze score, was used.

In the first experiment described here, text was prepared in which each psycholinguistic and SI oriented variable was manipulated individually in a controlled manner. Comprehension scores were analyzed separately to obtain added information on the effect of modification in level of single variables. This first experiment, accordingly, may be considered to represent a cross validation of the findings of Siegel and Bergman (1974) and of Lambert and Siegel (1974).

Experiment II was designed to assess the interactive effects of a set of SI oriented variables as they are systematically varied in level. A parallel investigation of interactive effects among selected psycholinguistic variables was completed in Experiment III.

In Experiment IV, passages were presented at exceptionally high and exceptionally low levels of CMU. The Experiment IV work provided some additional insight into the possible explanation for certain of the results emerging from Experiments II and III.

Levels of Variables

In Experiments I, II, and III, levels of manipulated variables were defined with reference to a set of descriptive norms developed by Williams, Siegel, Burkett, & Groff (in press). The norms were developed on 200 samples of technical material randomly taken from Air Force study guides, technical orders, career development course texts, and various manuals and regulations. Based on analysis of these samples, norms were developed which describe the level of each SI and psycholinguistically oriented variable found at each decile of the samples. "Low," "medium," and "high" levels of variables were defined in Experiments I, II, and III with reference to these decile values. A low level for any variable falls below the first decile value for that variable; a medium level is defined as falling between the fourth and sixth deciles, and a high level for any variable indicates that the variable fell above the ninth decile value. In all cases, variables were

defined in such a manner that high numerical values and high decile values were expected to be associated with higher comprehensibility and low numeric or decile values were expected to be associated with lower comprehensibility.

Comprehensibility

In all experiments, the criterion of comprehensibility employed was a cloze score. To obtain this score, text passages were modified by arbitrarily deleting words at the rate of 1 of 10 or 1 of 15 and replacing the deleted words with blank spaces of standard length. Individuals read the modified text and attempt to replace the deleted words. The proportion of correct replacements is called the cloze score. Since its development by Taylor (1953), the cloze technique has become a popular criterion measure of comprehensibility. The technique is held to be tolerant both to variation in the system of word deletion and to scoring strictness. Further, the technique is held to account for interest and prior knowledge of the reader (Taylor, 1953). Materials based on the cloze approach are easily constructed and do not require extensive pretesting or item validation (Klare, Sinaiko, & Stolurow, 1971).

Experiment I

Experiment I sought to investigate the effects of individually manipulating, within a textual passage, each of the psycholinguistic and SI oriented variables on cloze scores. As such, it served to verify and extend the work of Siegel and Bergman (1974) and of Lambert and Siegel (1974). The Structure-of-Intellect oriented variables involved were: CMU, CMR, MMU, ESI, NMS*, NMI, and DMU. The psycholinguistically oriented variables included were: YD, MD, TC, CE, LB, RB, and DC.

Textual Sample Selection

Forty-two passages were randomly selected from a set of current Air Force career development course (CDC) texts. Such texts are technical in nature and present the technical information required for an airman to advance in a career field in the Air Force. From two to five random selections were made from each volume of each of the following courses: CDC 43151C (Aircraft Maintenance Specialist, Jet Aircraft, One and Two Engines); CDC 43113 (Aircraft Mechanic); and CDC 64550 (Inventory Management Specialist). The samples taken were approximately 300 words in length. A sample was terminated at the first sentence end appearing after the 300th sampled word.

The selected passages were then randomly assigned to the 42 desired experimental conditions (14 variables x 3 levels of each = 42 conditions).

Preparation of Stimulus Materials

For the three blocks selected to represent a given variable, one block was rewritten at the low level for the variable; one block was rewritten at the medium level for the variable; and one block was rewritten at the high level for the variable. Thus, each of the

---

*NMS, convergent production of semantic systems, was included for investigation although, because of automatic measurement problems, this variable is not included in the CM program.

14 variables was tested in three different passages of structurally varied difficulty. No attempt was made to equate passage difficulty because the intent was to keep the passages as natural as possible and to test the effect of modifying a variable level on comprehensibility.

The value of the assigned variable was controlled within each of three subsections of 100 words into which each 300 word block was divided. This procedure ensured consistency throughout the passage and assured the accuracy of the value of the assigned variable.

Rewriting was performed with special emphasis on avoiding modifications not directly related to the relevant variable. Few, if any, sentence characteristic changes were made which were not basic to the measure of interest.

Following the preparation of the textual passages and verification of level of the assigned variable, cloze test forms were prepared by deleting every 10th word, the first word deleted in each passage being random selection from the first 10 words of the revised passage. Cloze test forms were typed in double spaced format. No modifications were made to subheads. Figures, diagrams, or tables referenced in the passages were not shown, since the variables under investigation are only measured in connected prose.

Test booklets were then assembled. Each booklet contained 21 passages--either numbers 1-21 (booklet A) or 22-42 (booklet B). Presentation order in each booklet was individually randomized. The variables assigned to each booklet are indicated in Table 4.

Table 4

Assignment of Variables to Booklets and Assignment of Subjects (Entry in Each Cell) by Reading Grade Level of Subjects

Level of Variable	LO				MED				HI				
Stimulus Booklet	A		B		A		B		A		B		
Subject Reading Level	LO	HI	LO	HI	LO	HI	LO	HI	LO	HI	LO	HI	TOTAL
Variable													
CMU	8	13			8	13					17	14	73
CMR			17	14			17	14	8	13			83
MMU			17	14	8	13			8	13			73
ESI	8	13					17	14	8	13			73
NMS			17	14	8	13			8	13			73
NMI	8	13					17	14			17	14	83
DMU	8	13			8	13					17	14	73
YD			17	14	8	13					17	14	83
MD			17	14	8	13					17	14	83
TC			17	14	8	13					17	14	83
SE			17	14	8	13					17	14	83
LB			17	14	8	13			8	12			72
RB			17	14	8	13			8	13			73
DC			17	14	8	13					17	14	83

Subjects

Two groups of subjects, one group representing moderate (high school level) reading ability and the other group representing high (college) reading ability, were involved. The moderate reading ability group included 25 paid volunteer subjects of a public vocational-technical high school. These subjects were enrolled in a variety of curricula, such as auto mechanics, electrical technology, food preparation, trowel trades, horticulture, and appliance repair. The high reading ability group was composed of 27 paid volunteer undergraduate college students. These subjects were majoring in a variety of fields, but all were then enrolled in an introductory psychology course. The assignment of subjects to passages is also indicated in Table 4.

Procedure

Data acquisition varied between two and one-half and four hours per subject based on individual work rate. The moderate reading level subjects completed all passages in a single session. The high reading level subjects completed their work in two sessions separated by one to two days. Because of the randomization procedures instituted and because of the rest periods allowed during the work, concern over possible fatigue effects seems unwarranted.

All subjects were initially administered the Nelson-Denny Reading Test (Revised), Form A (Brown, 1960). Instructions were then presented on the procedures for completing the cloze passages. The subjects were instructed to work at their own pace and told that both speed and accuracy were important. Breaks in the work were permitted. Booklets containing low numbered or high numbered passages were distributed to the subjects randomly. This was not a "speed test" in any sense of the term. All subjects completed all work and no time limit was imposed.

Results

Reading Level

As indicated in Figure 2, the moderate and the high reading level groups were well separated in terms of Nelson-Denny Reading Test Scores/RGL. The mean Nelson-Denny score of the moderate group was 48.6. This value is equivalent to an RGL of 10.00. The mean Nelson-Denny score of the high reading ability group was 85.7, which is equivalent to an RGL above 14.0, the highest point for which direct equivalents are reported. A Student's t test performed on these data indicates the difference to be statistically significant below the .01 level of confidence.

The Nelson-Denny Reading Test scores of members of the groups receiving the low and the high numbered passages within the moderate RGL and high RGL groups, and the overall subject group were compared. No evidence of differences was found in any comparison. The obtained t values were respectively: $t = .343$, $df = 23$; $t = .977$, $df = 25$; and $t = .339$, $df = 50$. Accordingly, the hypothesis of no difference in reading ability between the subgroups could not be rejected, and the data for the two subgroups were combined.

Cloze Test Scoring

Taylor (1953), in initially presenting the cloze procedure, reported that the procedure was not sensitive to the scoring criteria employed. According to Taylor, the resultant ordering of passage scores was not changed if synonyms of deleted words were accepted as correct or if only precise matches were accepted. In the current work, a relatively strict scoring procedure was followed:

- only precise matches of deleted words or matches involving errors of spelling, tense, or number were accepted as correct

- if a number was required, such as a table or figure number, model number, etc., any number, whether entered as digits or written out, was accepted.

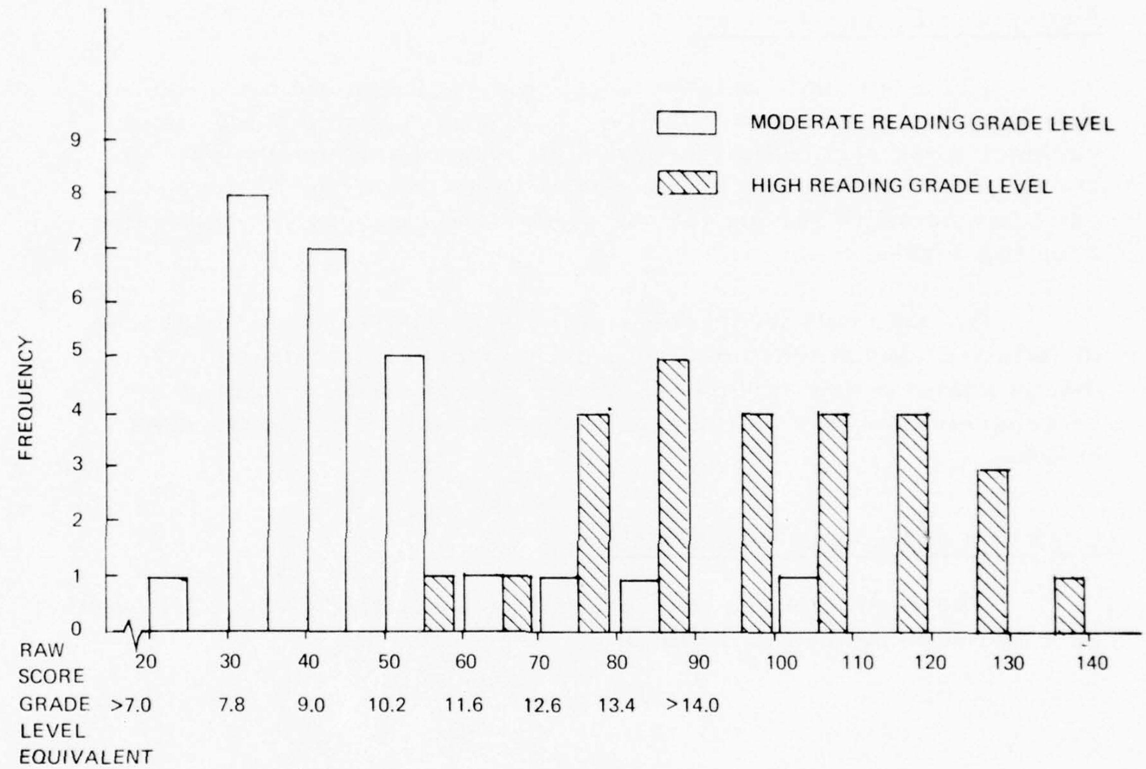


FIGURE 2. DISTRIBUTION OF NELSON DENNY RAW SCORES AND GRADE LEVEL EQUIVALENTS OF MODERATE AND HIGH READING GRADE LEVEL SUBJECT GROUPS IN EXPERIMENT I

In all cases, cloze test scores are reported as percentage of deleted words correctly entered, according to the criterion described above. This is the "normal" cloze score.

Analysis of Cloze Test Data

Cloze test data related to each psycholinguistic or SI variable were analyzed independently. Fourteen separate analyses of variance were accordingly performed. Each variance analysis investigated the effect on cloze score of variation of one of the 14 comprehensibility variables over three levels when RGL was varied over two levels.

Because cell frequencies in the analyses were not equal (due to failure of some scheduled subjects to appear), an unweighted-means analysis was performed (Winer, 1962). The results of the 14 separate analyses of variance are presented in the paragraphs below.

Cognition of Semantic Units (CMU)

The results of the variance analysis relative to the CMU variable are summarized in Table 5.

Table 5

Summary of Analysis of Variance for CMU Data

Source	SS	df	MS	F
CMU	.3067	2	.1534	19.67**
Reading Level	.1562	1	.1562	20.01**
CMU x Reading Level	.0112	2	.0056	< 1 n.s.
Within Cell	.5236	67	.0078	

** $p < .01$

Statistically significant differences were associated with both CMU level and reading group. These differences were statistically significant below the .01 level of confidence. The mean cloze score at each level of CMU for the moderate and the high reading groups is displayed graphically in Figure 3. The straight line of best fit for these data, as derived from the least square procedure, is also shown in Figure 3. Cloze score is seen to be positively sloped to a slight extent with increased CMU. The high reading ability group consistently scored higher than the low group.

Cognition of Semantic Relations (CMR)

Variation in level of the CMR variable also produced a statistically significant effect (Table 6) on cloze score. In this analysis, reading ability was again statistically significant at the .01 level of confidence. The mean data for this analysis are summarized in Figure 4, which indicates that cloze score increased substantially as CMR level was increased.

Table 6

Summary of Analysis of Variance for CMR Data

Source	SS	df	MS	F
CMR	.6708	2	.3354	19.85**
Reading Level	.1739	1	.1739	10.29**
CMR x Reading Level	.0234	2	.0117	< 1 n.s.
Within Cell	1.2975	77	.0169	

** $p < .01$

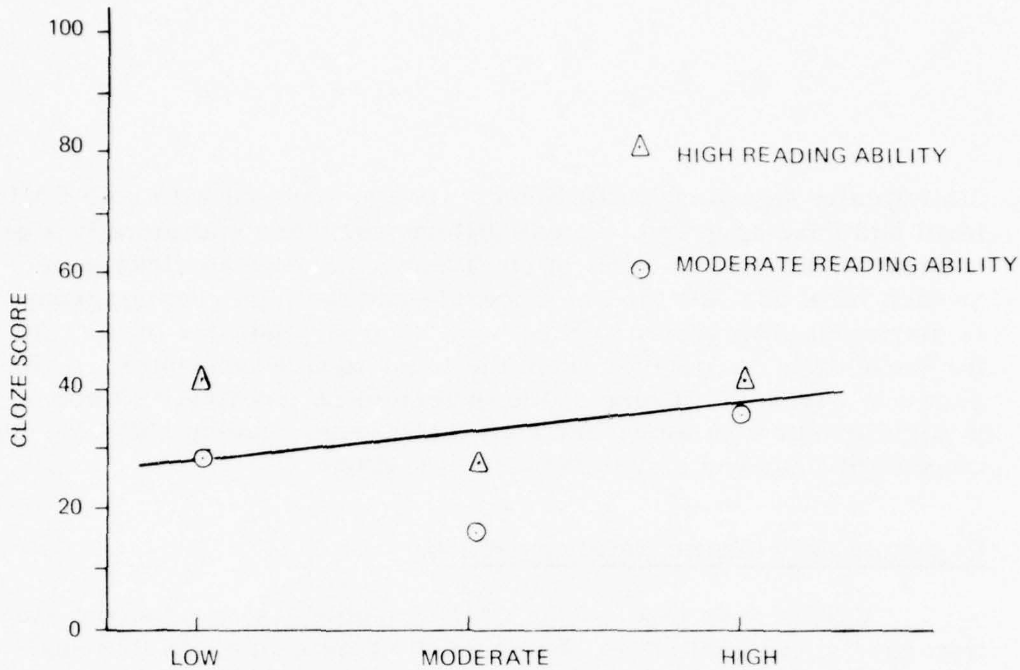


FIGURE 3. MEAN CLOZE SCORE AT THREE LEVELS OF CMU FOR MODERATE AND HIGH ABILITY READING GROUPS

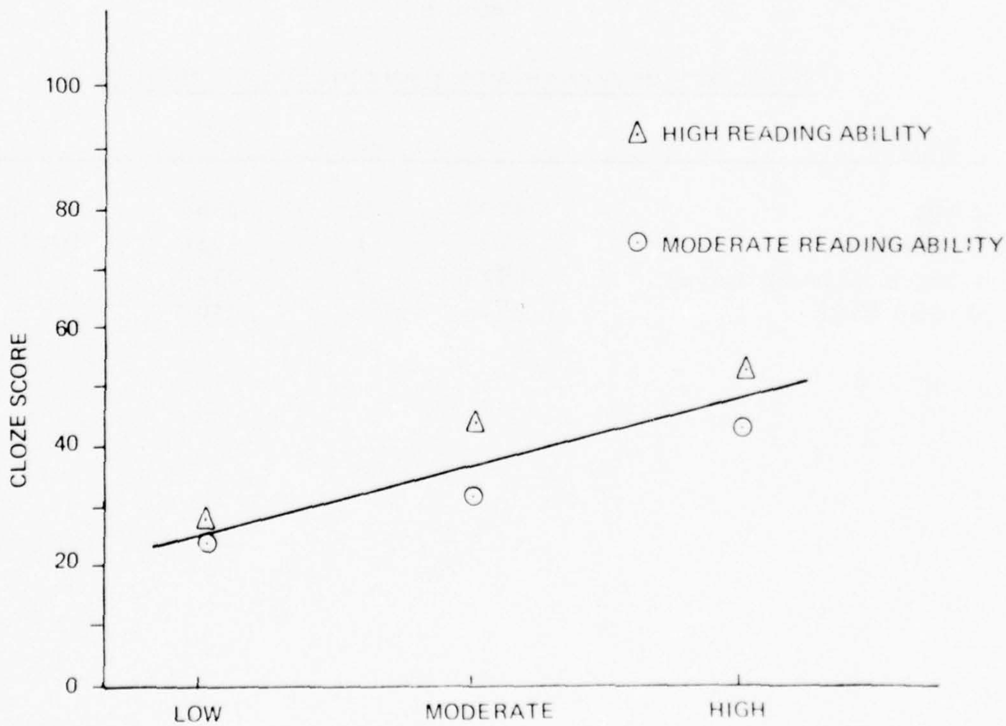


FIGURE 4. MEAN CLOZE SCORE AT THREE LEVELS OF CMR FOR MODERATE AND HIGH ABILITY READING GROUPS

Memory for Semantic Units (MMU)

As shown in Table 7, the MMU effects were statistically significant at the .01 level of confidence, as was the reading ability effect in this analysis. In Figure 5, the line of best fit slopes positively with increasing level of MMU. The high reading group mean scores were, again, greater than those of the moderate reading group.

Table 7

Summary of Analysis of Variance for MMU Data

Source	SS	df	MS	F
MMU	.8246	2	.4123	37.83**
Reading Level	.1741	1	.1741	15.97**
MMU x Reading Level	.0337	2	.0168	1.54 n. s.
Within Cell	.7332	67	.0109	

** $p < .01$

Evaluation of Symbolic Implications (ESI)

As shown in Figure 6, cloze score increased as a function of ESI level, and the means of high ability readers were higher than those of moderate ability. In the associated variance analysis, summarized in Table 8, both treatment effects were statistically significant below the .01 level of confidence.

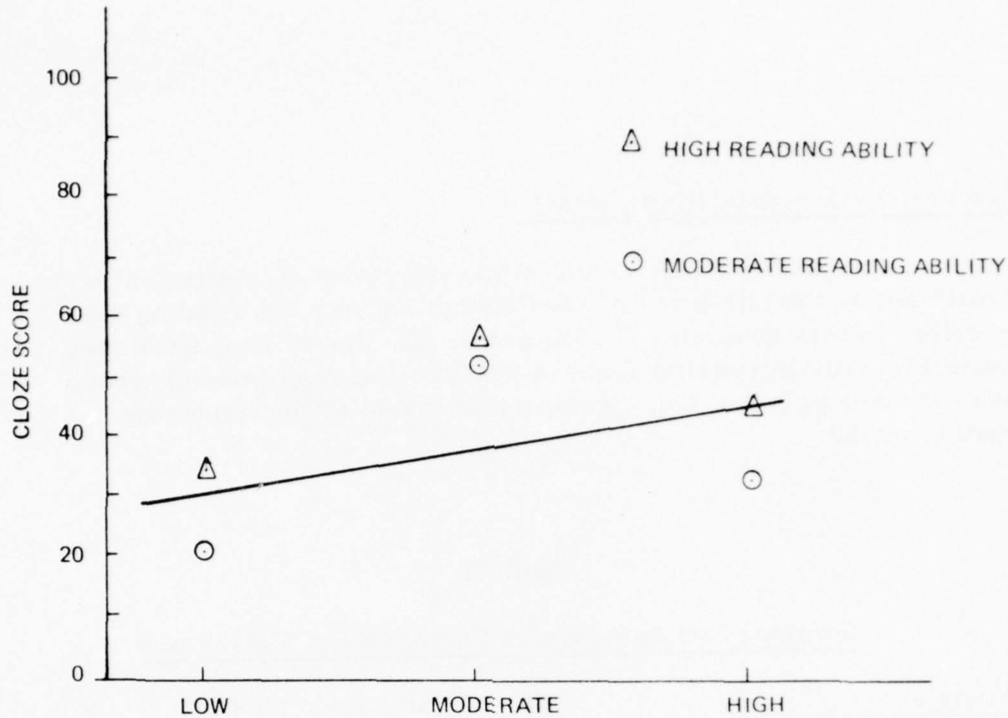


FIGURE 5. MEAN CLOZE SCORE AT THREE LEVELS OF MMU FOR MODERATE AND HIGH ABILITY READING GROUPS

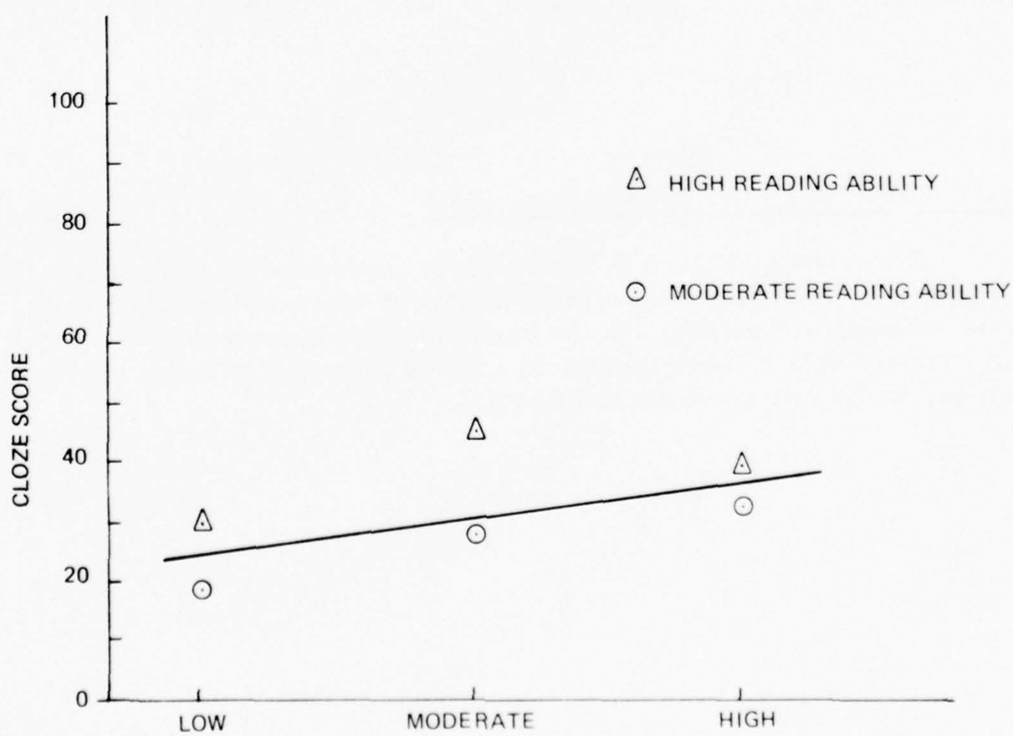


FIGURE 6. MEAN CLOZE SCORE AT THREE LEVELS OF ESI FOR MODERATE AND HIGH ABILITY READING GROUPS

Table 8

Summary of Analysis of Variance for ESI Data

Source	SS	df	MS	F
ESI	.2471	2	.1236	10.13**
Reading Level	.2449	1	.2449	20.07**
ESI x Reading Level	.0247	2	.0124	1.02 n.s.
Within Cell	.8207	67	.0122	

** $p < .01$

Convergent Production of Semantic Systems (NMS)

The NMS effect and reading level were statistically significant at the .01 level of confidence in this fifth analysis (summarized in Table 9). Figure 7 indicates that cloze scores were lower at high levels of NMS, and that high ability readers, on the average, out-scored moderate ability readers. The data trend was not in the anticipated direction and is discussed in the subsequent discussion section relative to the set of SI variables. We note here, however, that NMS is not included in the CM program.

Table 9

Summary of Analysis of Variance for NMS Data

Source	SS	df	MS	F
NMS	.4853	2	.2426	26.66**
Reading Level	.2022	1	.2022	22.22**
NMS x Reading Level	.0213	2	.0106	1.16 n.s.
Within Cell	.6095	67	.0091	

** $p < .01$

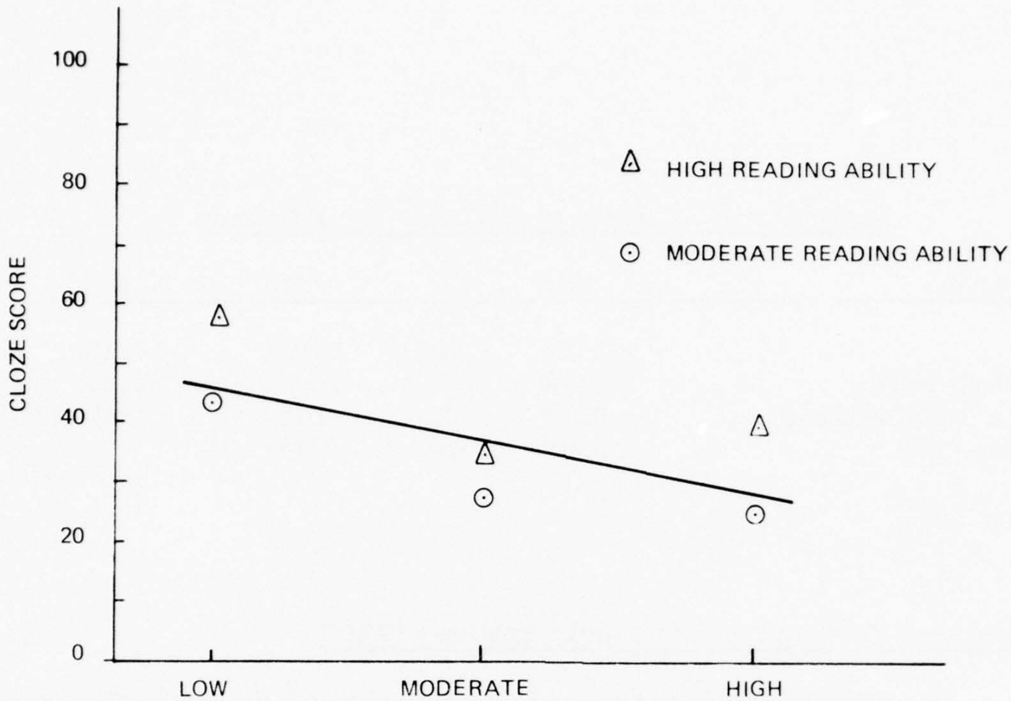


FIGURE 7. MEAN CLOZE SCORE AT THREE LEVELS OF NMS FOR MODERATE AND HIGH ABILITY READING GROUPS

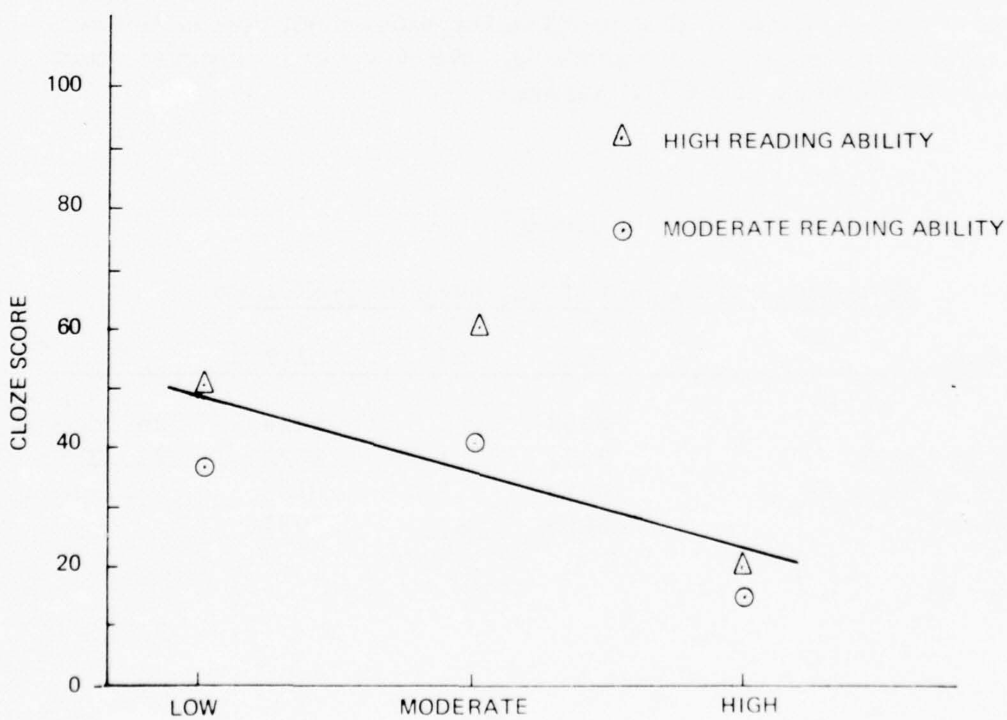


FIGURE 8. MEAN CLOZE SCORE AT THREE LEVELS OF NMI FOR MODERATE AND HIGH ABILITY READING GROUPS

Convergent Production of Semantic Implications (NMI)

As for NMS, cloze score decreased with increases in NMI. Again, the high ability readers outscored the moderate ability readers (Figure 8). These were again statistically significant. The summary of the variance analysis of these data is presented in Table 10. The negative slope is discussed in the discussion section for the Structure-of-Intellect variables.

Table 10

Summary of Analysis of Variance for NMI Data

Source	SS	df	MS	F
NMI	1.5116	2	.7558	55.17**
Reading Level	.3127	1	.3127	22.82**
NMI x Reading Level	.0675	2	.0338	2.47 n.s.
Within Cell	1.0554	77	.0137	

** $p < .01$

Divergent Production of Semantic Units (DMU)

Variation in DMU level, as shown in the variance analysis summary of Table 11, produced effects which were statistically significant below the .01 level of confidence. The reading level effect in this analysis similarly was statistically significant below the .01 level. Figure 9 shows that mean cloze score was positively related to DMU value.

Table 11

Summary of Analysis of Variance for DMU Data

Source	SS	df	MS	F
DMU	.7167	2	.3584	31.72**
Reading Level	.1595	1	.1595	14.12**
DMU x Reading Level	.0045	2	.0022	< 1 n.s.
Within Cell	.7588	67	.0113	

** $p < .01$

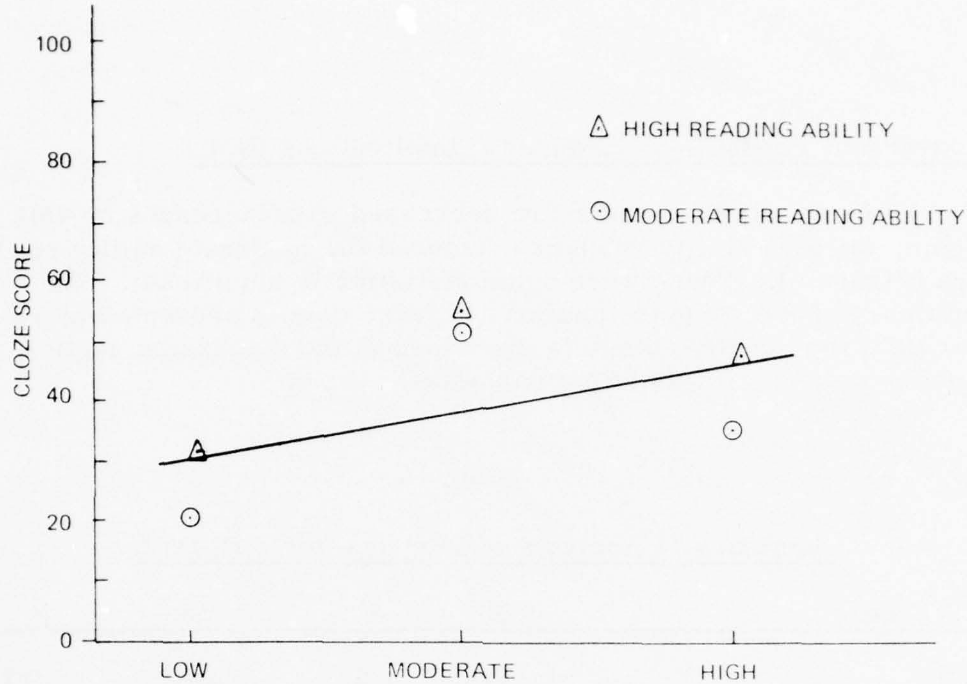


FIGURE 9. MEAN CLOZE SCORE AT THREE LEVELS OF DMU FOR MODERATE AND HIGH ABILITY READING GROUPS

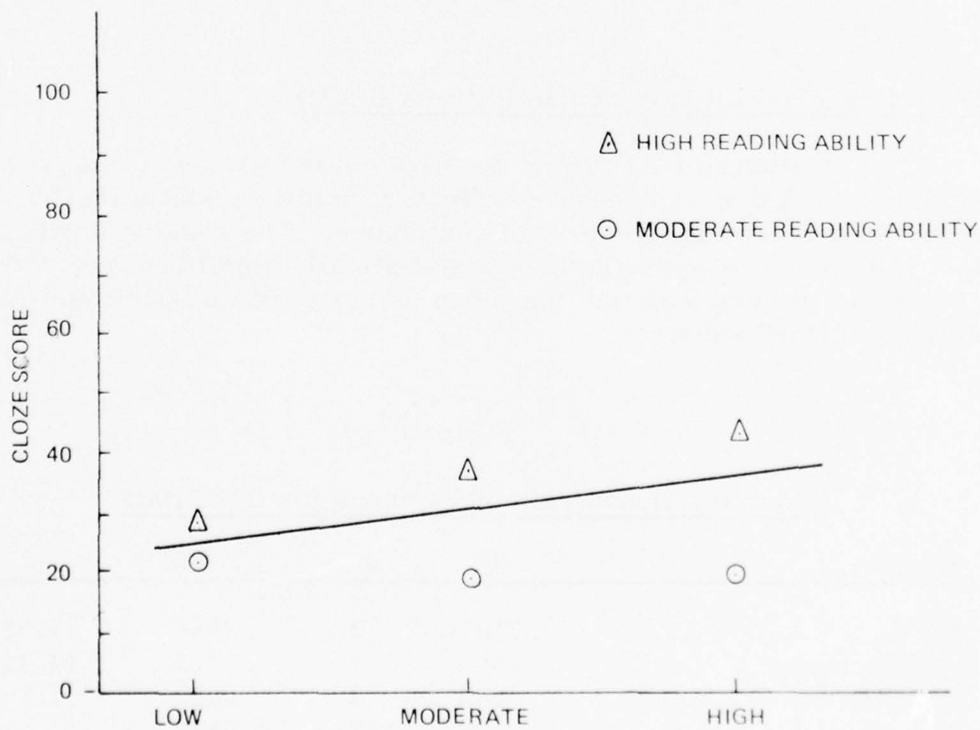


FIGURE 10. MEAN CLOZE SCORE AT THREE LEVELS OF YD FOR MODERATE AND HIGH ABILITY READING GROUPS

Yngve Depth (YD)

Variation in level of YD, a psycholinguistically oriented variable, produced effects which were statistically significant below the .01 level of confidence. Higher mean levels of the YD measure were associated with elevated cloze scores (Figure 10) and higher reading ability persons tended to achieve higher mean cloze scores. This difference was also statistically significant at the .01 level of confidence.

Table 12

Summary of Analysis of Variance for YD Data

Source	SS	df	MS	F
YD	.1778	2	.0889	10.96**
Reading Level	.1570	1	.1570	19.38**
YD x Reading Level	.0272	2	.0136	1.68 n.s.
Within Cell	.6243	77	.0081	

** $p < .01$

Morpheme Depth (MD)

According to the analysis summarized in Table 13, the MD level variation produced an effect on cloze score which was statistically significant at the .01 level of confidence. The interaction with reading ability was statistically significant at the same level. The high reading ability subjects consistently produced higher cloze scores than the moderate ability readers. Figure 11 presents the obtained trend.

Table 13

Summary of Analysis of Variance for MD Data

Source	SS	df	MS	F
MD	.4126	2	.2063	19.28**
Reading Level	.2089	1	.2089	19.52**
MD x Reading Level	.0882	2	.0441	4.12 n.s.
Within Cell	.8273	77	.0107	

** $p < .01$

Transformational Complexity (TC)

While variation in level of TC affected cloze score at or below the .01 level of confidence, the mean data points (Figure 12) for each reading ability group essentially formed a symmetrical "V." The straight line of best fit was essentially horizontal. Again, the *reading grade level effect was statistically significant.*

Table 14

Summary of Analysis of Variance for TC Data

Source	SS	df	MS	F
TC	.3776	2	.1888	14.52**
Reading Level	.6565	1	.6565	50.50**
TC x Reading Level	.0234	2	.0117	< 1 n.s.
Within Cell	.9974	77	.0130	

** $p < .01$

Center Embeddedness (CE)

The analysis of variance summarized in Table 15 indicates that level of CE significantly affected cloze score at or below the .01 level of confidence, as did reading ability. The trends for the line

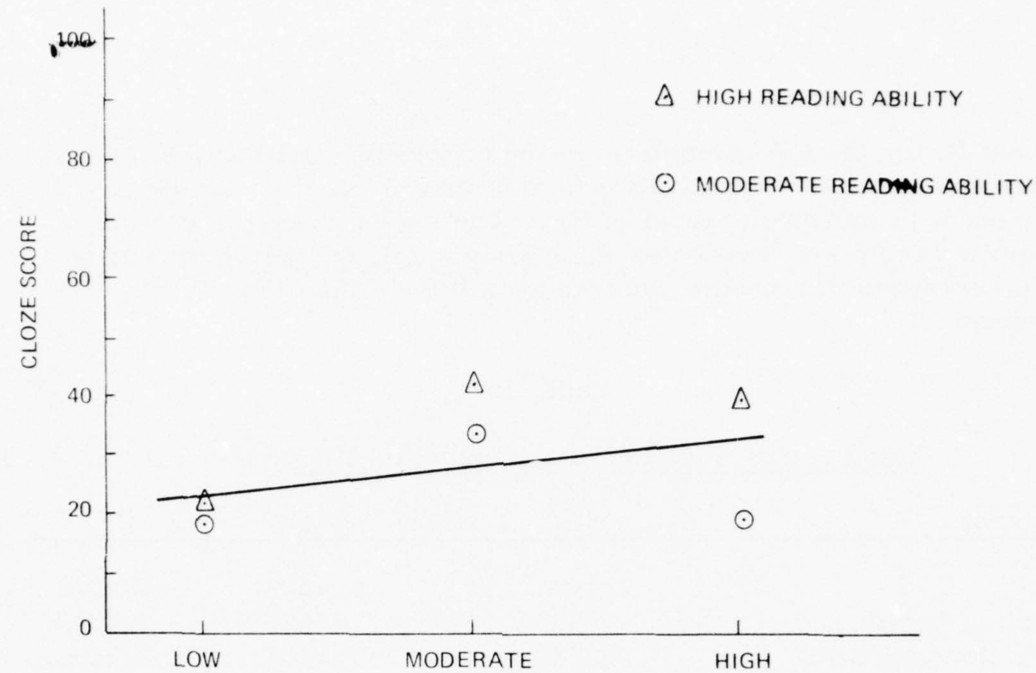


FIGURE 11. MEAN CLOZE SCORE AT THREE LEVELS OF MD FOR MODERATE AND HIGH ABILITY READING GROUPS

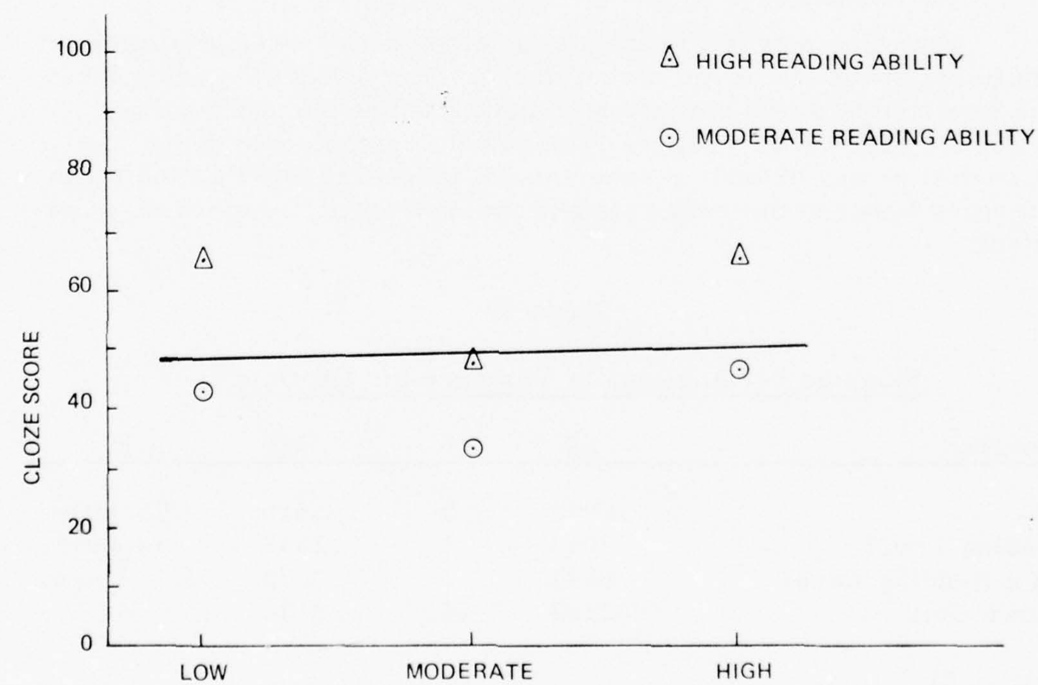


FIGURE 12. MEAN CLOZE SCORE AT THREE LEVELS OF TC FOR MODERATE AND HIGH ABILITY READING GROUPS

of best fit for the CE variable is in the nonpredicted direction. This is discussed later. Figure 13 indicates that mean cloze score decreased with increasing level of CE. The subjects appear to have benefitted from low levels of CE, and high ability readers appear to have benefitted to a greater degree at this level than low ability readers.

Table 15

Summary of Analysis of Variance for CE Data

Source	SS	df	MS	F
CE	.9176	2	.4588	35.29**
Reading Level	.1142	1	.1142	8.78**
CE x Reading Level	.0973	2	.0486	3.74 n.s.
Within Cell	1.0038	77	.0130	

** $p < .01$

Left Branching (LB)

The effects of left branching on cloze score were statistically significant at the .01 level of confidence. The effect of reading level in this analysis was statistically significant at the .01 level of confidence (Table 16). Figure 14 shows that mean cloze score increased in a very orderly manner as LB increased and that the mean difference between the moderate and the high ability readers was consistent.

Table 16

Summary of Analysis of Variance for LB Data

Source	SS	df	MS	F
LB	.7038	2	.3519	25.50**
Reading Level	.2043	1	.2043	14.80**
LB x Reading Level	.0144	2	.0072	< 1 n.s.
Within Cell	.9102	66	.0138	

** $p < .01$

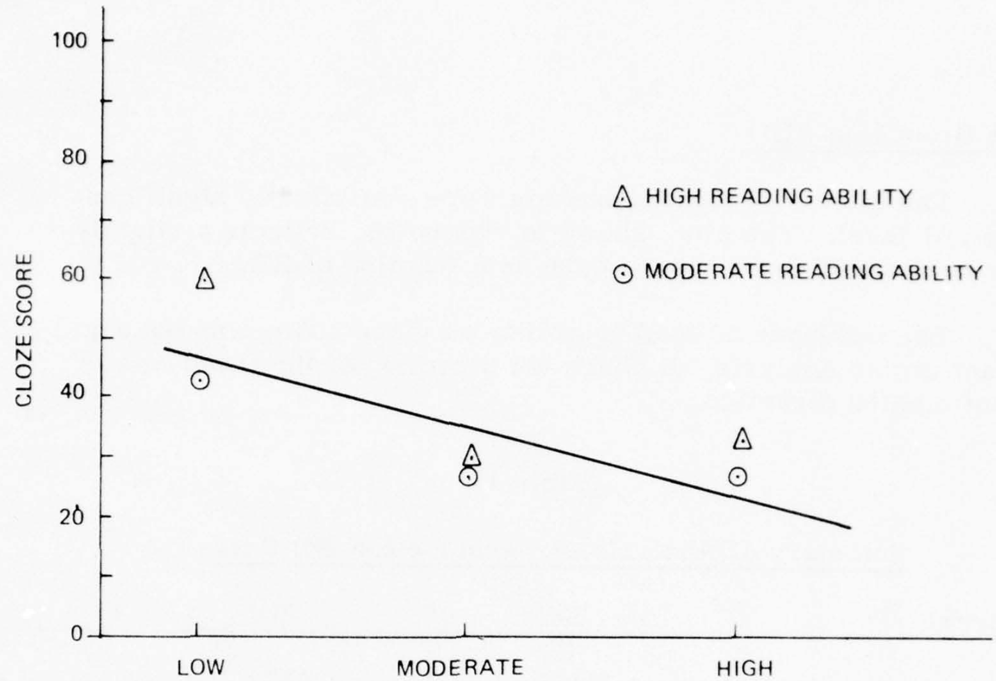


FIGURE 13. MEAN CLOZE SCORE AT THREE LEVELS OF CE FOR MODERATE AND HIGH ABILITY READING GROUPS

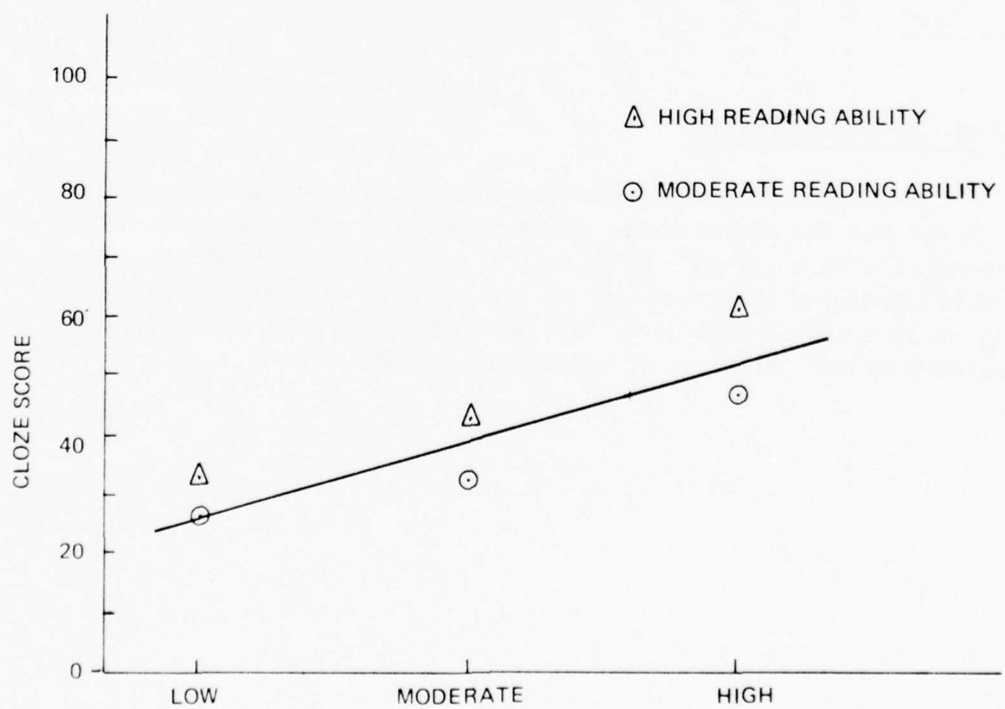


FIGURE 14. MEAN CLOZE SCORE AT THREE LEVELS OF LB FOR MODERATE AND HIGH ABILITY READING GROUPS

Right Branching (RB)

The effects of right branching were statistically significant at the .01 level. The plot, shown in Figure 15, reflects a slightly negative slope of mean cloze score as a function of RB.

The influence of reading ability on cloze score was not significant in this analysis, although the direction of the trend was in the anticipated direction.

Table 17

Summary of Analysis of Variance for RB Data

Source	SS	df	MS	F
RB	.1146	2	.0573	5.67**
Reading Level	.0006	1	.0006	< 1 n. s.
RB x Reading Level	.0101	2	.0050	< 1 n. s.
Within Cell	.6736	67	.0101	

** $p < .01$

Deleted Complement (DC)

Complement deletion affected cloze score at the .01 level of confidence and the slope of the mean cloze score data as a function of the level of this variable (Figure 16) was slightly negative. Again, the high reading ability subjects on the average outscored the moderate reading ability subjects, and the difference was statistically significant at the .01 level of confidence (Table 18).

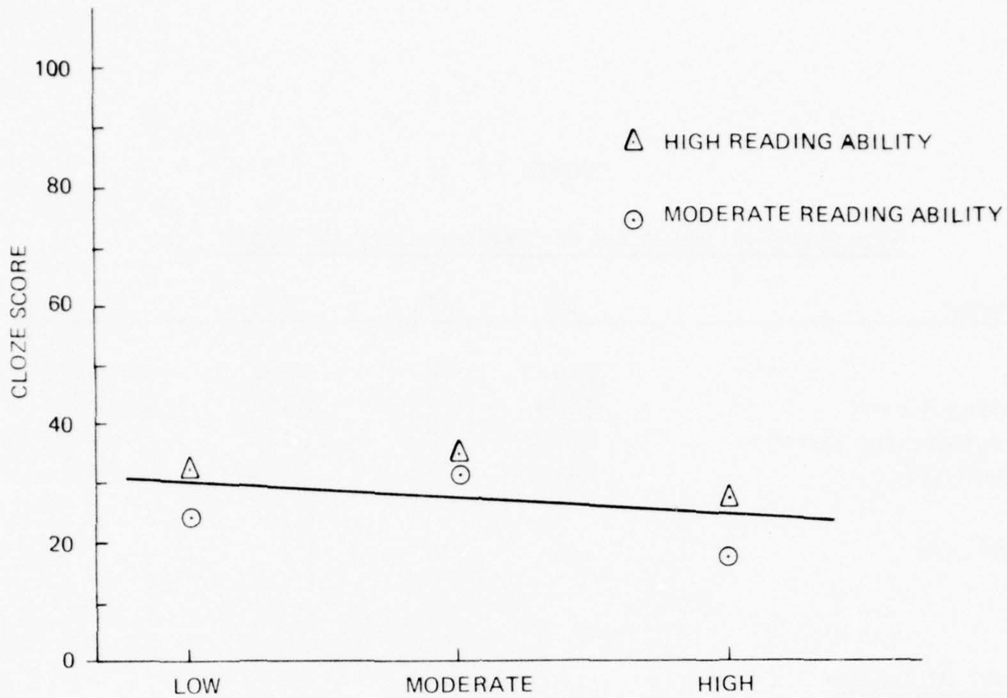


FIGURE 15. MEAN CLOZE SCORE AT THREE LEVELS OF RB FOR MODERATE AND HIGH ABILITY READING GROUPS

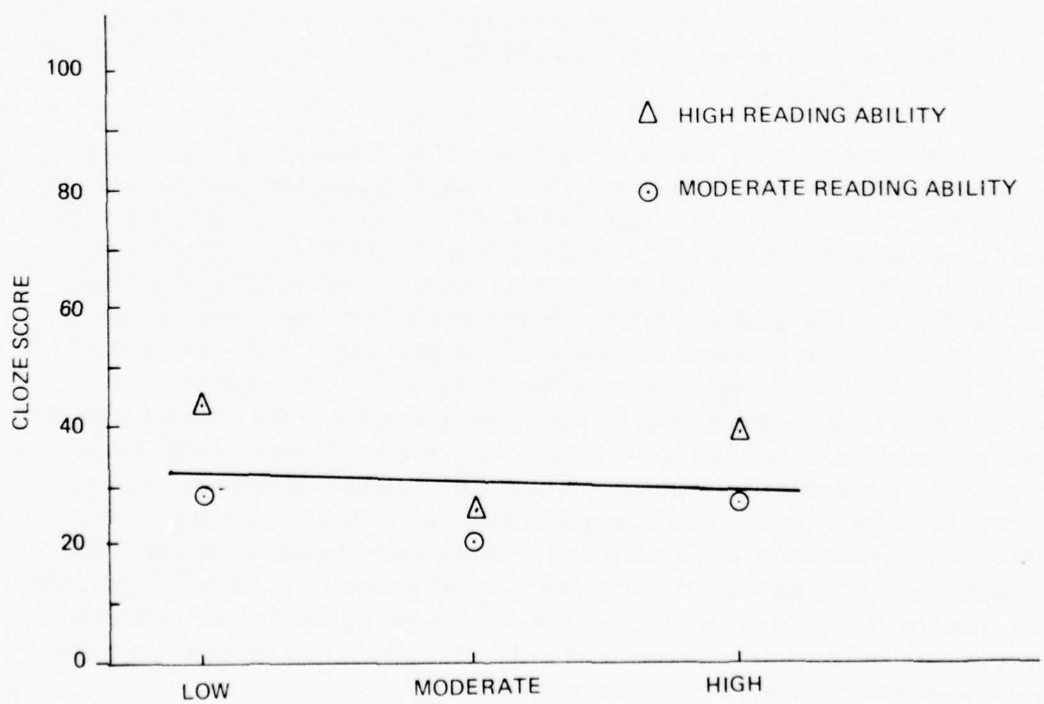


FIGURE 16. MEAN CLOZE SCORE AT THREE LEVELS OF DC FOR MODERATE AND HIGH ABILITY READING GROUPS

Table 18

Summary of Analysis of Variance for DC Data

Source	SS	df	MS	F
DC	.2504	2	.1252	10.18**
Reading Level	.2504	1	.2504	20.36**
DC x Reading Level	.0272	2	.0136	1.11 n.s.
Within Cell	.9477	77	.0123	

** $p < .01$

Trend Analyses

The several analyses of variance were extended in order to determine the absence or presence of significant quadratic trends across the examined levels of the studied variables. Since the interactions between variable and reading level were not statistically significant in any of the cases, the analyses were continued as single factor analyses. Techniques appropriate to unweighted means analyses were continued.

In the cases of three variables--YD, CMR, and LB--linear trends reached significance ($p < .01$), while quadratic trends did not. In seven analyses--MMU, ESI, NMS, NMI, DMU, MD, and CE--both the linear and the quadratic trends attained statistical significance below the .01 level. In the remaining four cases--CMU, TC, RB, and DC--only the quadratic trends achieved statistical significance ($p < .01$). These findings indicate variables YD, CMR, and LB to be linear in their relationship to the criterion. The comprehensibility variables in the group of seven displayed a linear trend which was augmented by a quadratic trend, and the variables in the final group of four exhibited a purely quadratic trend. However, as pointed out by Hicks (1964), a quadratic trend must be viewed with caution in situations such as that involved here because a quadratic equation may be passed through any three points. In order to verify the quadratic trends, the variables should be examined at additional levels. It is for this reason that only straight lines of best fit have been presented in Figures 3 through 16.

Discussion

Structure-of-Intellect Variable Results

In a concurrent effort, the same SI and psycholinguistic variables as considered here were employed as a basis for developing a multiple linear regression equation relating cloze score to variable score (Williams, Siegel, Burkett, & Groff, in press). Comparison of the present results with the results indicated by this regression work provided one basis for evaluating the present indications. A second evaluative basis rested on an integration of the present results with those of the Siegel and Bergman (1974) study in which the same SI variables were involved. The results of the investigation of the SI variables in the current work and in the other mentioned efforts are summarized in Table 19. As indicated by Table 19, significant effects and positive correlations between level of intellectual load (assumed to be imposed by levels of each variable) and comprehension consistently support the SI variables as measures of textual comprehensibility, with the exception of NMS (convergent production of semantic systems) and NMI (convergent production of semantic implications).

Table 19

Summary of Structure-of-Intellect Results in Current Work and Other Efforts

	Siegel & Bergman		Williams et al.		Current Work	
Variable	Probability	Point Biserial r	Probability	Product Moment r	Probability	R
CMU	< .025	.28	< .01	.39	< .01	.56
CMR	< .002	.56	< .01	.19	< .01	.55
MMU	< .005	.35	< .01	.38	< .01	.68
ESI	< .001	.31	< .01	.33	< .01	.43
NMS	< .005	.33	< .01	.18	< .01	.61
NMI	< .001	.39	< .01	-.11	< .01	.72
DMU	< .001	.38	n. s.	-.01	< .01	.56

There are a number of possible reasons for the differences between the current results and those of the related prior work with regard to NMI. The NMI variable evolved considerably since the exploratory work of Siegel and Bergman in which it was defined with regard to degree of completeness of syllogisms contained within sample passages. In the current study, the variable was defined with regard to the mean number of parts of speech of words contained within a passage. Although both approaches were believed to measure the same construct, the possibility of sensitivity differences exists.

Aside from definitional differences, there were a number of other differences between the two studies. Siegel and Bergman employed a much wider range of NMI than was employed in the present work. Additionally, they used Air Force enlisted personnel as subjects, as compared with the technical vocational high school and college subjects of the present study. It is believed that some similarity in reading ability, interests, and mental ability exists among the various groups, but the possibility of subject differences remains open. Additionally, the Siegel and Bergman textual materials were not of the technical training nature, as were the textual materials of the present work.

The correlations resulting from the regression equation work and which are included in Table 19 are correlations between variable scores and cloze scores for the complete subject group employed in the multiple regression work. Separate correlational values, based on data provided by low ability readers and by high ability readers (as defined by Nelson-Denny Reading Test Scores) were also available as the result of the regression work. The product moment correlation between NMI variable score and cloze score for low ability readers was -.007. The correlation, based on high ability reader data was -.214, or essentially zero. Accordingly, the negative slope of comprehension as a function of NMI is consistent with other results relating level of NMI, as currently defined, and cloze score.

NMS, convergent production of semantic systems, has been consistently defined throughout. Accordingly, differences between the present and the prior work cannot be attributed to definitional problems. However, the other differences between the Siegel and Bergman work and the present work hold and may be causative. In this regard, we note that an extreme restriction of range was imposed on the present NMS variable range resulting from the norms used to identify decile levels. The low, medium, and high text passages had NMS values respectively of: 0.000, 0.062, and 0.375, equivalent to means of 0.00, 0.28, and 1.50 comprehension aids per hundred words of text.

Finally, the reading group by SI interactions were remarkably free from statistical significance. This result serves to extend the potential of the findings because it suggests that variable effectiveness was not differential across reading groups.

Psycholinguistic Variable Results

The present data, relevant to psycholinguistically oriented variables, are discussed here in relationship to the regression work, the results of the exploratory work of Lambert and Siegel (1974), and the results obtained by other workers studying the same language variables. The findings in the current analyses and correlations between variable level and cloze score obtained during the regression work are summarized in Table 20.

Lambert and Siegel (1974) reported Yngve depth to affect inconsistently comprehensibility in their own work and in that of others. The effects of this factor appeared to be partly a function of the method by which YD was examined. Lambert and Siegel concluded that while the variable may be method sensitive, it may be useful in determining the mental load imposed by a written passage. In the current analysis, a slight positive slope was found in the functional relationship between cloze score and YD. Similarly, a positive correlation coefficient was obtained for the corresponding correlation in the development of the readability regression equation. Accordingly, there is continued support for considering YD as a factor affecting comprehensibility.

Table 20

Summary of Psycholinguistic Results in Current Work and Related Efforts

Variable	Current Work Probability	Current Work R	Williams et al. Probability	Williams et al. Product Moment r
YD	< .01	.57	< .01	.13
MD	< .01	.66	< .01	.33
TC	< .01	.46	< .01	.25
CE	< .01	.85	n. s.	-.02
LB	< .01	.78	n. s.	-.04
RB	< .01	.89	< .01	-.16
DC	< .01	.56	n. s.	-.02

Lambert and Siegel found evidence that MD influenced comprehension when text was examined at the sentence level and at the paragraph level. Surveyed literature also consistently reported the same relationship. In the development of the comprehensibility regression equation, a positive correlation (.33) was found between MD level and cloze score. In the current work, a significant effect due to MD was found, and the slope of the trend line was positive. Examination of Figure 11 shows that the mean moderate reading ability subject cloze score was not highest at the high MD level. Nonetheless, the variable remains viable as a comprehensibility variable.

Transformational complexity has been investigated as a determinant of readability many times. Studies reported by Lambert and Siegel (Coleman, 1964, 1965; Slobin, 1966; Wason, 1959; 1961; etc.), as well as more recent work (Evans, 1972-73; Peltz, 1973-74) support this concept as a comprehensibility variable. Lambert and Siegel's work was consistent with this trend, as were findings of the current work and those of the readability equation development.

Lambert and Siegel provided moderate agreement with the findings of Schwartz et al. (1970) and of Wang (1970), that center embeddedness interfered with comprehension. In the current work, a significant effect was associated with this variable. However, the direction of the trends in both the current work and in the regression work were opposite to those hypothesized. In these cases, the dependent measure was cloze score, unlike those of the other cited studies. The apparently consistent negative trend may be evidence of method bias with regard to this variable.

Left branching and right branching seem to affect comprehension inconsistently, judged by findings of Schwartz et al. (1970), Hamilton and Deese (1971), and Lambert and Siegel (1974). The current results supported left branching and right branching as variables for predicting the comprehensibility of written materials. The regression results were inconsistent with the present results, suggesting these to be weak variables.

Complement deletion was hypothesized by Lambert and Siegel to degrade comprehension, following Fodor and Garrett (1967) and Hakes (1972). Lambert and Siegel's data did not support their hypothesis and indicated, in fact, a trend in the opposite direction. The current results show that deletion of complements aided comprehension to a slight degree, and the regression equation data do not negate these results. We note that the rarity of occurrence of complements or their deletions should preclude the finding, in the practical situation, of a powerful effect due to this variable. Current theory describes complements as markers of sentence structure. Repeated findings that such markers interfere with comprehension would argue that current conceptions regarding complements are incomplete.

Finally, we note that a negative slope or a horizontal trend does not negate the value of a comprehensibility variable. Clearly, such trends might indicate only that the effect of the variable diminishes along the scale or that the effect of the modification was relatively constant.

Experiment II

The data for Experiment II were actually acquired prior to Experiment I. The second experiment reported here sought evidence of interactive effects among SI variables relative to comprehensibility. To this end, each of three SI oriented measures was varied within a factorial design. The SI variables were selected on several bases: freedom from apparent redundancy with other measures, minimal overlap of underlying Guilford categories, range of numerical values of the variable as measured during norm development, and inclusion in the Comprehensibility Measurement (CM) computer program in its form at the time of the study. The chosen variables were CMU, ESI, and NMI, as indicated in Table 21.

Table 21

Comparison of Structure-of-Intellect Variables

Variable	Normative Range	Comments	Chosen Variables
CMU	.293 - .476	Two intellective categories common with CMR; vocabulary oriented	X
CMR	.008 - .070	Two intellective categories common with CMU; vocabulary oriented	
MMU	.783 - .874	One intellective category common with CMU and CMR; vocabulary oriented	
ESI	.943 - 1.000	Not vocabulary oriented as are CMU, CMR, and MMU	X
NMS	0.000 - .375	Not in computer program	
NMI	.388 - .605	Reasoning	X
DMU	0.000 - .175	Reasoning	

Various details of Experiment II were treated as they were in Experiment I. Levels of manipulated variables were defined with respect to the set of norms developed by Williams et al. (in press) exactly as they were in Experiment I. The dependent variable employed in Experiment II was, again, cloze score.

The experimental paradigm is presented in Figure 17.

	CMUH*			CMUM*			CMUL*		
	NMIH	NMIM	NMIL	NMIH	NMIM	NMIL	NMIH	NMIM	NMIL
ESIH	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
ESIM	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
ESIL	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-

(REPEAT FOR LOW RGL SUBJECTS)

*H= high
M= moderate
L= low

Figure 17. Paradigm for Experiment II.

Textual Sample Selection

Twenty-seven passages were randomly selected from CDC texts supplied for the purpose by the Air Force. From one to three random selections were made from each text volume of each of three courses: CDC 43151C (Aircraft Maintenance Specialist, Jet Aircraft, One and Two Engine); CDC 43113 (Aircraft Mechanic); and AFSC 64559 (Inventory Management Specialist). The samples taken were generally 135 to 145 words in length and a sample was terminated after the first sentence end appearing after the 135th sampled word. In a few cases, this rule produced slightly longer samples. In these instances, if the final sentence of the sample was a compound sentence, the sample was terminated at a junction of adjacent independent clauses occurring after 135 or more words.

Preparation of Stimulus Materials

All combinations of low, medium, and high level of the three selected SI oriented measures were required for the experiment as designed. One sample CDC passage was randomly assigned to each of the 27 treatment conditions. The level of the CMU, ESI, and NMI variables was then measured in each sample. Each sample passage was then revised if necessary so that the level of the three measured variables fell within the range assigned for that respective sample.

It was possible in this sort of rewriting to achieve desired levels through manipulations such as making separate sentences of complex sentences, blindly replacing words with synonyms to influence vocabulary diversity, and the like. Manipulations such as these were not allowed in the revisions. Modifications were required to affect the variable of interest in the equations, and modified materials were required to appear similar in tone and style to the original passages.

The RGL of the originally selected CDC samples was measured through application of the Automated Readability Index of Smith and Senter (1966). RGL's ranged from 7.76 to 16.51, with the median RGL of the sample at 11.50. In order to avoid confounding obtained cloze scores by variation in textual sample reading difficulty level, all revised samples were modified as necessary to cause all measured RGL's to fall between 11.25 and 11.75. The equation of Smith and Senter employs word length in letters and

sentence length in words for prediction of RGL. These factors were not involved in the measures under study. Hence, revision of sample passages for the purpose of modifying RGL should minimally affect the independent measures studied.

The passages were set in cloze test form by deleting every tenth word following a word deleted at random from the first ten words. A packet of the 27 cloze test forms with presentation order individually randomized was prepared for each experimental subject.

Test Pacing

A group of five high school students completed a portion of the cloze tests to permit estimation of proper pacing of the test passages. Time required for individuals to complete various passages ranged from 8 minutes to 14 minutes. In order to impose a moderate time stress on the subsequent subjects, a time allotment of 8 minutes was selected for work on a single passage. Accordingly, the results obtained may be limited to the situation in which a slight time stress is involved. We note, however, that the situation was not rushed. All subjects finished all the work.

Subjects

Sixteen paid volunteer subjects participated. Eight were solicited from a local college and eight were respondents to a classified advertisement seeking individuals who had terminated their formal education during or prior to the tenth grade.

Procedure

Data collection was performed during a seven hour period of a single day. First, the Nelson-Denny Reading Test (Revised) Form A (Brown, 1960) was administered to all subjects. They were then instructed in the procedures to be followed for completing the cloze forms. Eight minutes were allowed for each sample, and the subjects were warned when one minute of working time remained. A ten minute break was taken after every six passages. Fifteen passages in addition to the Nelson-Denny Reading Test were completed in the morning. The remaining 12 were completed after a 60 minute lunch break. The order of passage presentation was randomized. Accordingly, fatigue effects are believed to be equally distributed over the data.

Reading Level Results

As in Experiment I, the mean reading ability levels of the group of subjects were well separated. As indicated in Figure 18, there was no overlap in the distributions of Nelson-Denny test scores of the two groups. The mean raw score of the low reading ability group was 43.8 (equivalent to the 9.5 grade level) and the mean raw score of the high reading ability group was 95.6 (corresponding to a reading grade level above the 14th grade). A t test indicated this difference to be statistically significant below the .001 level of confidence ($t = 5.82$, $df = 14$). Accordingly, the two groups can be considered to represent separate populations.

Cloze Test Scoring

The cloze form scoring for Experiment II was identical in all respects to the scoring procedures followed in Experiment I.

Data Treatment

The data of Experiment II were analyzed through a variance analysis. The results are summarized in Table 22.

Main Effects

Of the SI oriented variables, only the effects of ESI were statistically significant. Plots of the main effects are shown in Figures 19, 20, and 21. The plots indicated that CMU was associated with cloze score in the expected way. ESI variation affected cloze score of low ability readers in the expected way, but its effects on cloze scores for high ability readers as well as the effects of NMI on cloze for both reading ability groups were irregular.

The large, consistent difference between the cloze scores obtained by high and by low ability readers was statistically significant at or below the .01 level of confidence.

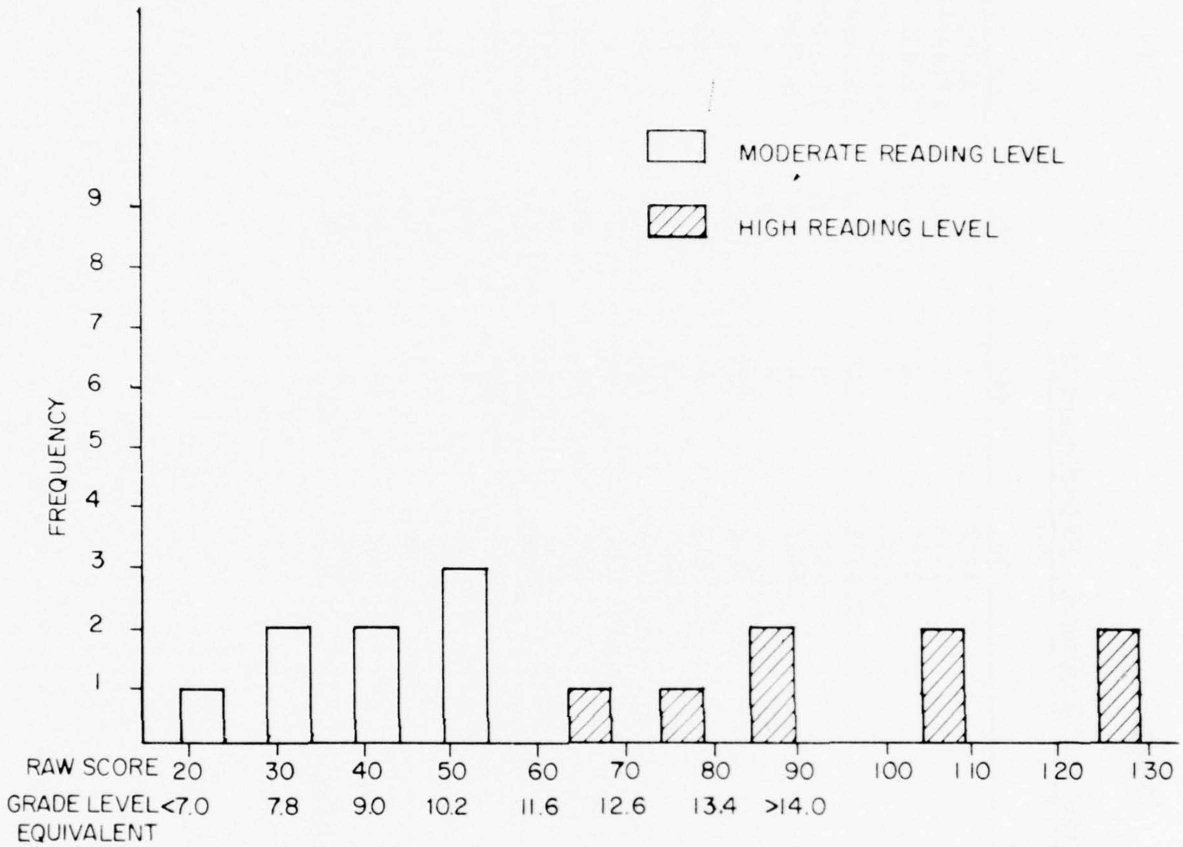


FIGURE 18. DISTRIBUTIONS OF NELSON DENNY RAW SCORES AND GRADE LEVEL EQUIVALENTS OF MODERATE AND HIGH READING LEVEL SUBJECT GROUPS IN EXPERIMENT II.

Table 22

Summary of Analysis of Variance, Experiment II

	SS	*df*	*MS*	*F*
CMU	.0548	2	.0274	3.262
ESI	.3919	2	.1960	15.835**
NMI	.0580	2	.0290	2.302
GRADE LEVEL	4.6854	1	4.6854	33.515**
SUBJ: GRADE LEVEL	1.9578	14	.1398	13.573**
CMU x ESI	.5847	4	.1462	11.246**
CMU x NMI	.3862	4	.0967	6.810**
CMU x GRADE LEVEL	.0104	2	.0052	< 1
CMU x SUBJ W/IN GRADE LEVEL	.2345	28	.0084	< 1
ESI x NMI	.1610	4	.0402	4.232**
ESI x GRADE LEVEL	.0084	2	.0042	< 1
ESI x SUBJ W/IN GRADE LEVEL	.3450	28	.0123	1.194
NMI x GRADE LEVEL	.0037	2	.0018	< 1
NMI x SUBJ W/IN GRADE LEVEL	.3537	28	.0126	1.223
CMU x ESI x NMI	.6044	8	.0756	7.340**
CMU x ESI x GRADE LEVEL	.1159	4	.0290	2.231
CMU x ESI x SUBJ W/IN GRADE LEVEL	.7274	56	.0130	12.621**
CMU x NMI x GRADE LEVEL	.0386	4	.0097	< 1
CMU x NMI x SUBJ W/IN GRADE LEVEL	.7924	56	.0142	1.379
ESI x NMI x GRADE LEVEL	.0436	4	.0109	1.147
ESI x NMI x GRADE LEVEL	.5324	56	.0095	< 1
CMU x ESI x NMI x GRADE LEVEL	.1948	8	.0244	2.369*
(RESIDUAL) CMU x ESI x NMI x SUBJ W/IN GRADE LEVEL	1.1560	112	.0103	< 1

* $p < .05$
** $p < .01$

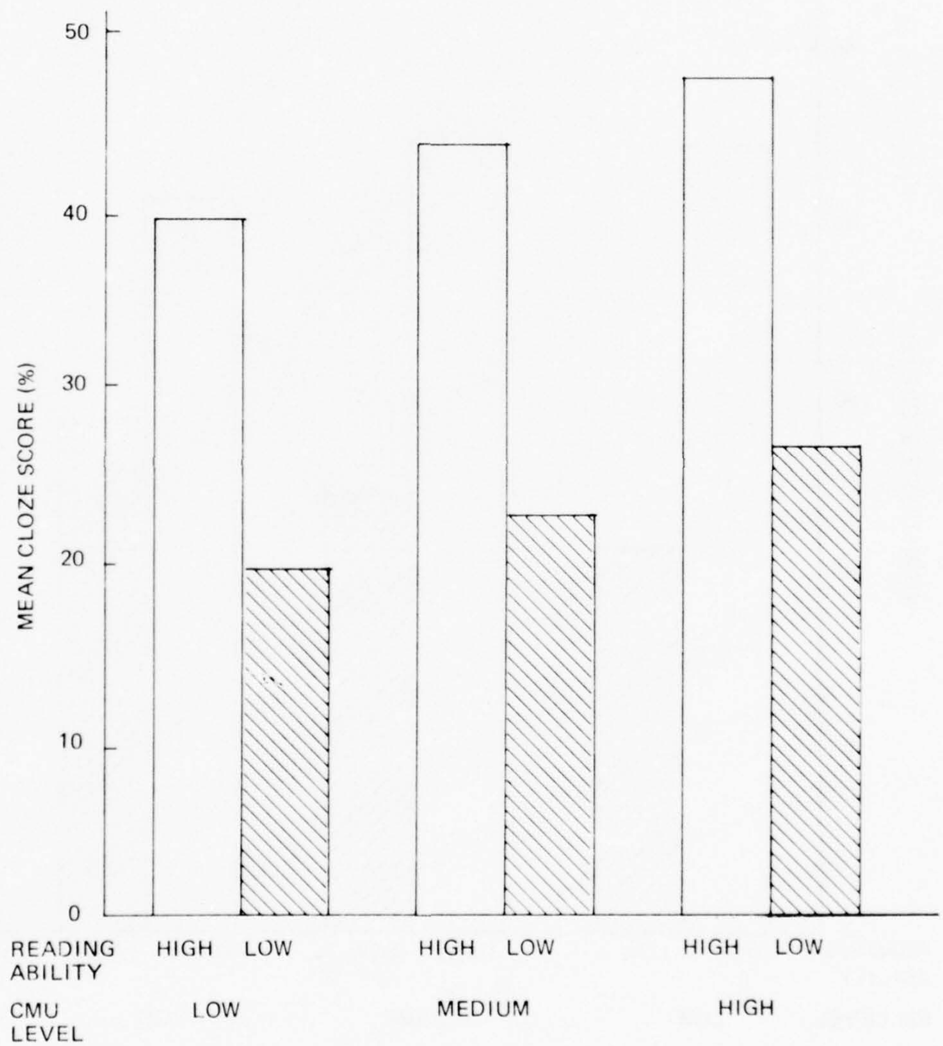


FIGURE 19. MEAN CLOZE SCORE FOR HIGH AND LOW READING ABILITY SUBJECTS AT THREE LEVELS OF CMU

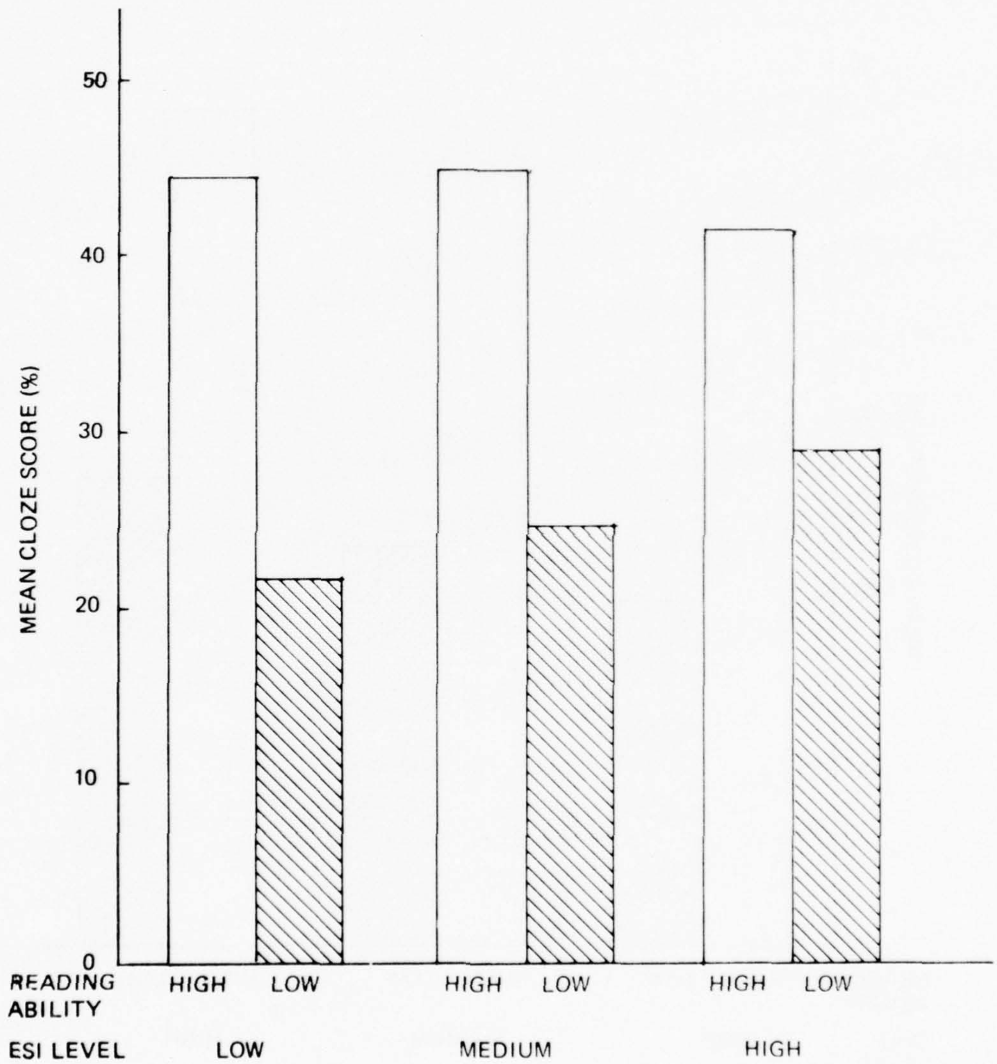


FIGURE 20. MEAN CLOZE SCORE FOR HIGH AND LOW READING ABILITY SUBJECTS AT THREE LEVELS OF ESI

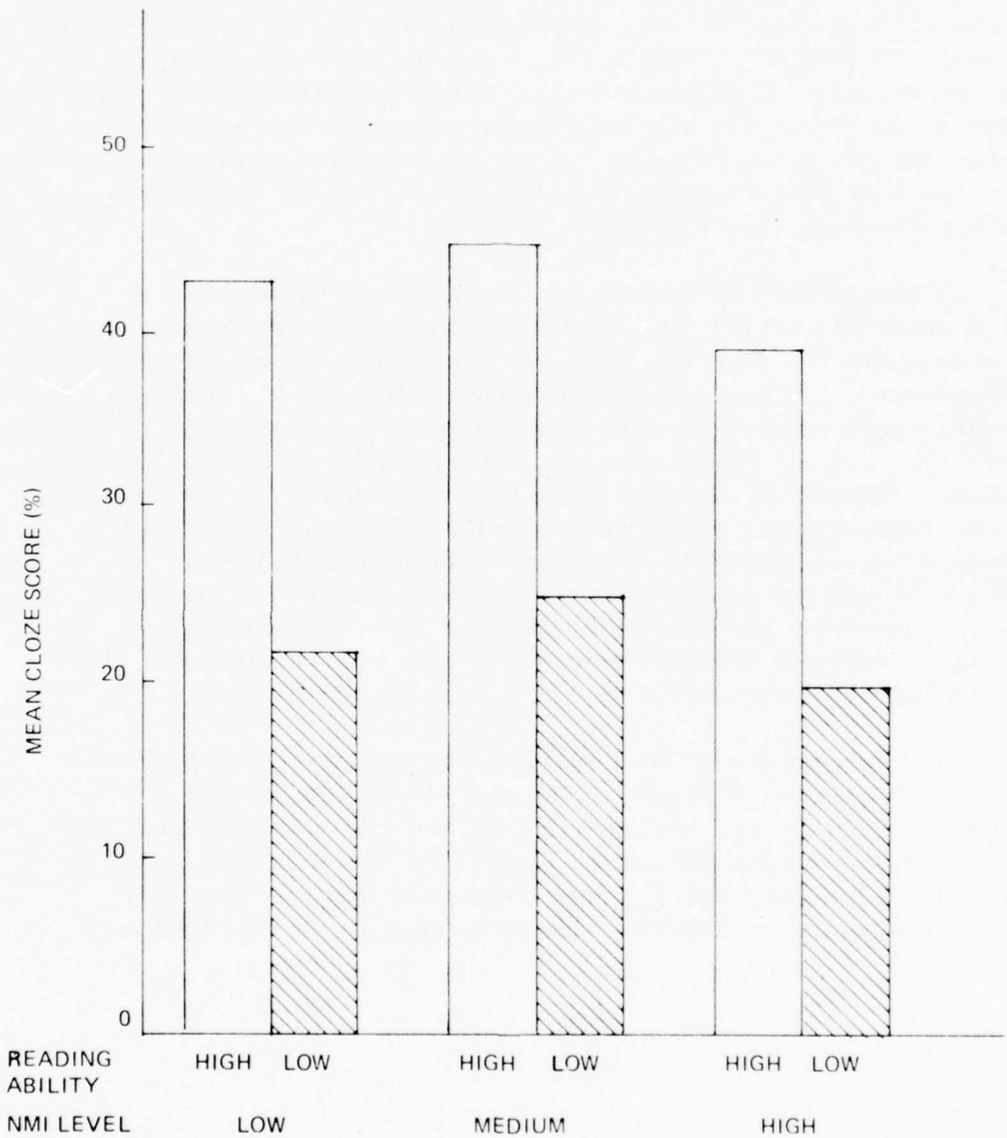


FIGURE 21. MEAN CLOZE SCORE FOR HIGH AND LOW READING ABILITY SUBJECTS AT THREE LEVELS OF NMI

Interactions

Only the two-way interactions for the Structure-of-Intellect variables are discussed. The statistically significant CMU and ESI interaction is plotted in Figure 22. At medium to high levels of CMU, increasing ESI appears to have raised comprehension. At the low to medium level of CMU, modification of ESI produced irregular changes in cloze score. It appears that removal of abbreviations (raising ESI) improves comprehension, unless extreme vocabulary diversity (low CMU) is present.

The statistically significant interaction between CMU and NMI is plotted in Figure 23. Both variables, although very different in method of measurement, were based on vocabulary oriented measures. They may be somewhat related because the short, common words which are most frequent in language, and which raise CMU, are also the words of multiple parts of speech, lowering NMI. This would seem to make the two additive (noninteractive). However, the additive trend was not fully indicated. Specifically, the high NMI-medium CMU point was considerably below the expected level and the low CMU curve was almost horizontal. Apparently, lowering the number of parts of speech was effective when the number of different words was high but not when the number of different words was low. This seems entirely logical.

The interaction of NMI and ESI was statistically significant at the .01 level of confidence. The plot of this interaction is presented in Figure 24 and, as anticipated, the plot indicates that the interaction was relatively weak. Generally, a linear trend is indicated by these data with a strong indication that medium NMI was more affected by ESI than either the high or the low NMI levels.

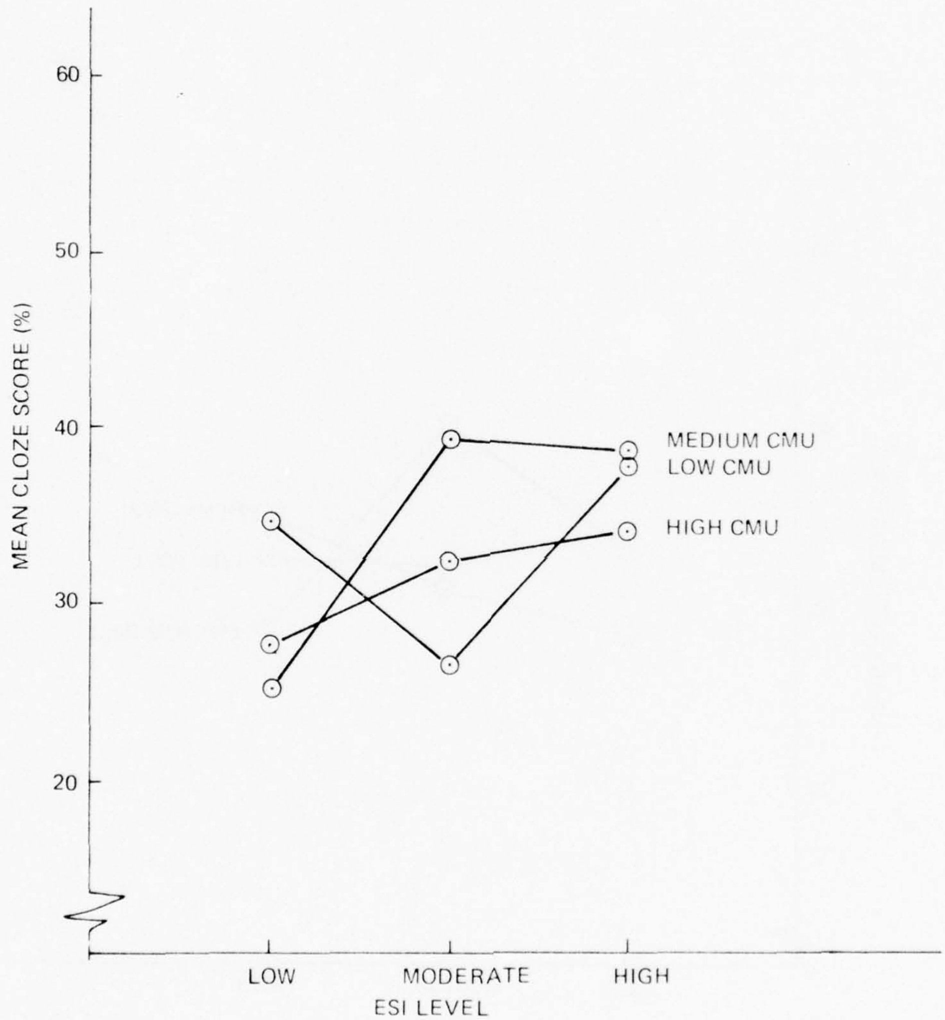


FIGURE 22. INTERACTION OF CMU AND ESI

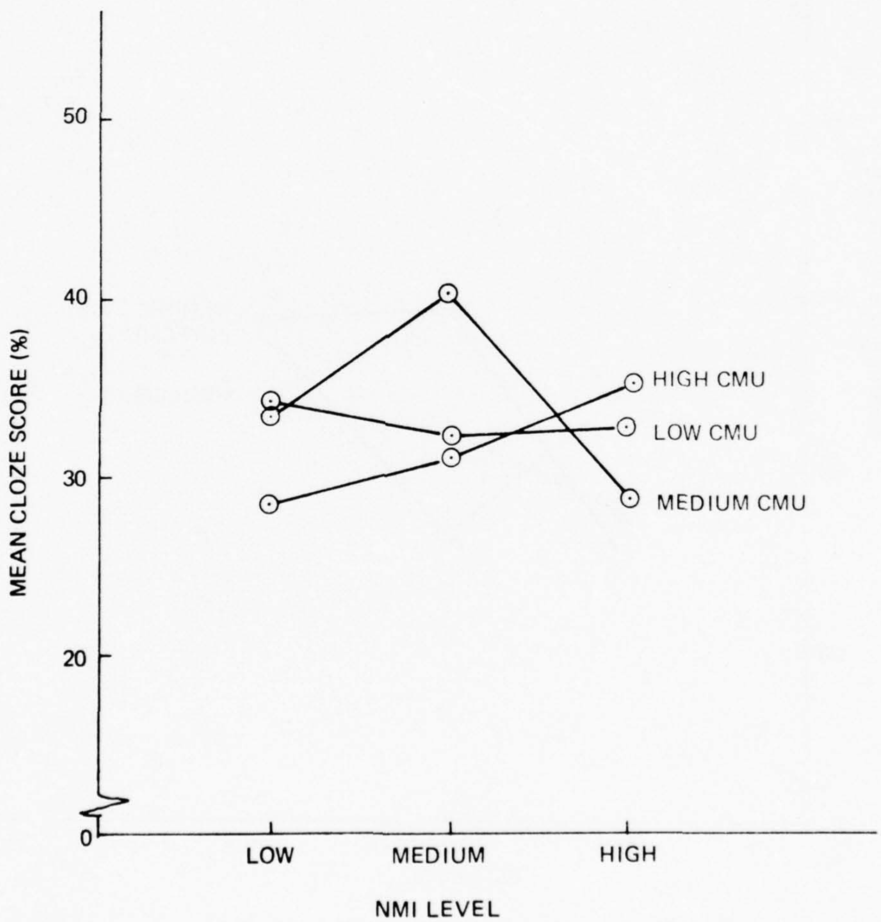


FIGURE 23. INTERACTION OF CMU AND NMI

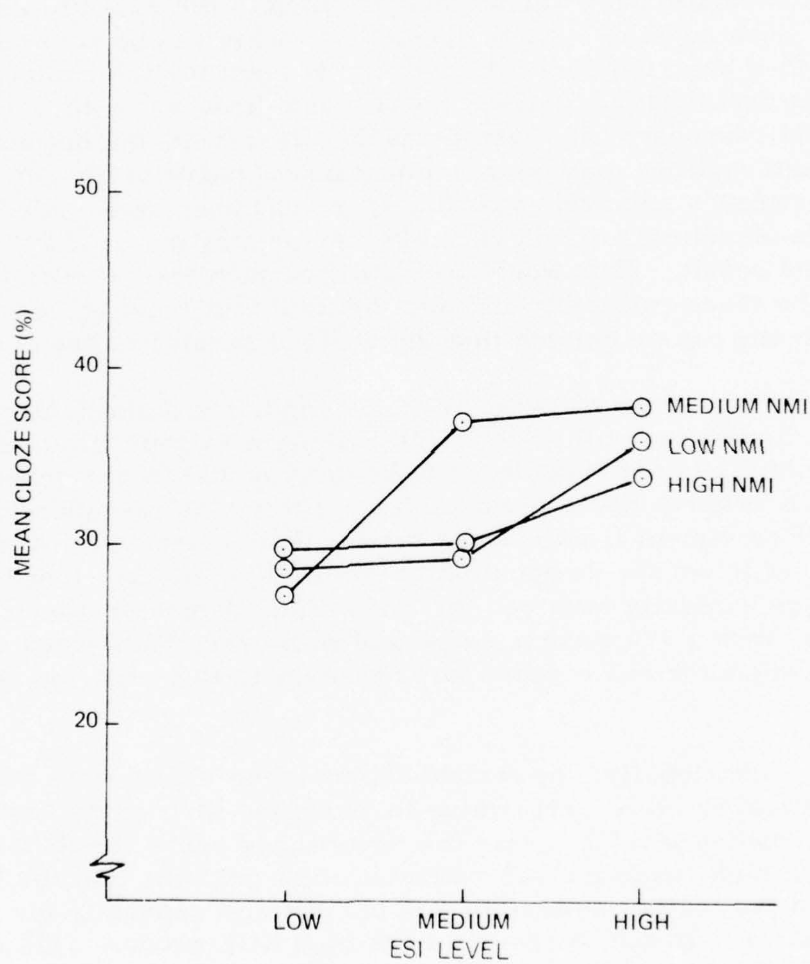


FIGURE 24. INTERACTION OF ESI AND NMI

Discussion

The major purpose of Experiment II was to investigate the possibility of interactive effects among three selected SI oriented variables purported to measure textual comprehensibility. Such interactivity was evidenced and was interpreted as indicating a complex relationship among the variables in the real life situation. Such complex effects are not considered to be unusual nor were they unanticipated. Moreover, at least for the comprehensibility measures employed, the two and three way interactions seemed reasonable and interpretable. However, the presence of such interactions may have implications relative to the rewriting of text after a comprehensibility score has been determined. Simple adjustment of one variable may or may not produce the desired result. This would indicate that such textual adjustments must be made rationally and that the total text must be reevaluated before one can be certain that the desired result has been obtained.

The failure of Experiment II to indicate statistically significant findings for all main effects, along with some main effect irregularity is not considered to be detrimental to our general posture relative to the SI variables. The textual samples employed in Experiment II were quite brief. While such small samples were sufficient for demonstrating interactive effects, longer samples are evidently required for main effect demonstration. For example, with a 135 word passage and with every tenth word deleted, each variable's cloze score was based on 13 to 14 fill-ins (135/10 = 13.5).

Additionally, the textual stimulus materials were set at a RGL equal to 11.5. This value is above the RGL of the best moderate reading level subjects (Figure 18) and below the RGL of the poorest high reading level subjects. It is possible that the 11.5 RGL of the textual materials was too difficult generally for the low RGL group and too easy for the high RGL group. This may have tended to mask main effect sensitivity.

Experiment III

Experiment III paralleled Experiment II in nature and scope, but sought to obtain evidence of interaction effects among psycholinguistically oriented textual measures. The criteria for selecting the included psycholinguistically oriented variables, experimental design, and data collection procedure were nearly identical to those of Experiment II. Only differences from the procedures of Experiment II will be described here.

The variables chosen for investigation in Experiment III were selected for minimal redundancy, breadth of normative range, and inclusion in the automated textual analytic computer program. The selected variables were: YD (an overall measure of structural complexity), MD (a vocabulary oriented measure), and LB (which is based on the existence of particular characteristics of sentences as parsed for determination of Yngve depth). A comparison of the psycholinguistically oriented variables is presented as Table 23.

Table 23

Comparison of Psycholinguistic Variables on Experiment III Variable Selection Criteria

Variable	Normative Range	Comments	Chosen Variables
YD	.508 - .676	Overall sentence complexity	X
MD	.568 - .742	Vocabulary measure	X
TC	.917 - .998	Portion of sentence complexity	
SE	.123 - 1.000	Portion of sentence complexity	
LB	0.000 - 1.000	Portion of sentence complexity	X
RB	.220 - .405	Not significant, Experiment I	
DC	.892 - 1.000	Occurs rarely	

Textual Sample Selection

The 27 test passages selected for use in Experiment II provided the basis for the test passages prepared for Experiment III.

Preparation of Stimulus Materials

The passages were randomly assigned to experimental levels and modified as required The policies followed for ensuring preparation of valid passages in Experiment II were also followed in Experiment III. Finished samples fell between grade levels 11.25 and 11.75, as measured by the Automated Readability Index, as was the case in Experiment II.

The cloze test forms were prepared as in Experiment II. The order of test forms within packets was individually randomized. The procedures of Experiment II were again followed.

Test Pacing

The time allowance for Experiment II passage completion was found to be unnecessarily long. In Experiment III, five minutes were allowed for each passage. The subjects were warned when one minute remained in the time allotted for each passage. Again, all subjects completed all passages within the time allowance and the randomization procedures were assumed to distribute fatigue effects equally over the data.

Subjects

Eight college students (high reading ability) and eight students in a technical job training program (moderate ability students) served as subjects. All subjects were paid, as before.

Procedure

The design of the study paralleled that of Experiment II. Data collection required three hours. The high reading ability students were tested in two separate sessions of one and one-half hours each

on successive days. The moderate reading ability subjects were tested in a single session. A ten minute break was taken every hour.

As before, the data collection sessions began with administration of the Nelson-Denny Reading Test, Form A (Brown, 1960).

Reading Level Results

The reading level of members of the subject groups in the current experiment are presented in Figure 25. As in Experiment II, the groups were well separated in reading level. The mean raw score of the moderate reading level group was 31.0. This value is equivalent to grade level 7.9. The mean score of the high ability group was 107.0. This value is above the 14th grade level. The difference between the groups was statistically significant below the .001 level of confidence ($t = 7.85$, $df = 14$).

Cloze Test Scoring

The cloze test scoring was completed in the same manner as for Experiment II.

Data Treatment

As before, variance analysis constituted the principal data analytic tool. The results of the analysis of the variance of the data are presented in Table 24.

Main Effects

For the main effects (YD, MD, and LB), only the effects of LB exerted a statistically significant influence. The main effect mean data are presented in Figures 26, 27, and 28. Figure 28 indicates that the LB result was due to variation of cloze scores of high ability readers as the level of LB was varied. There was little, if any, variation across levels of other variables for either reading ability group.

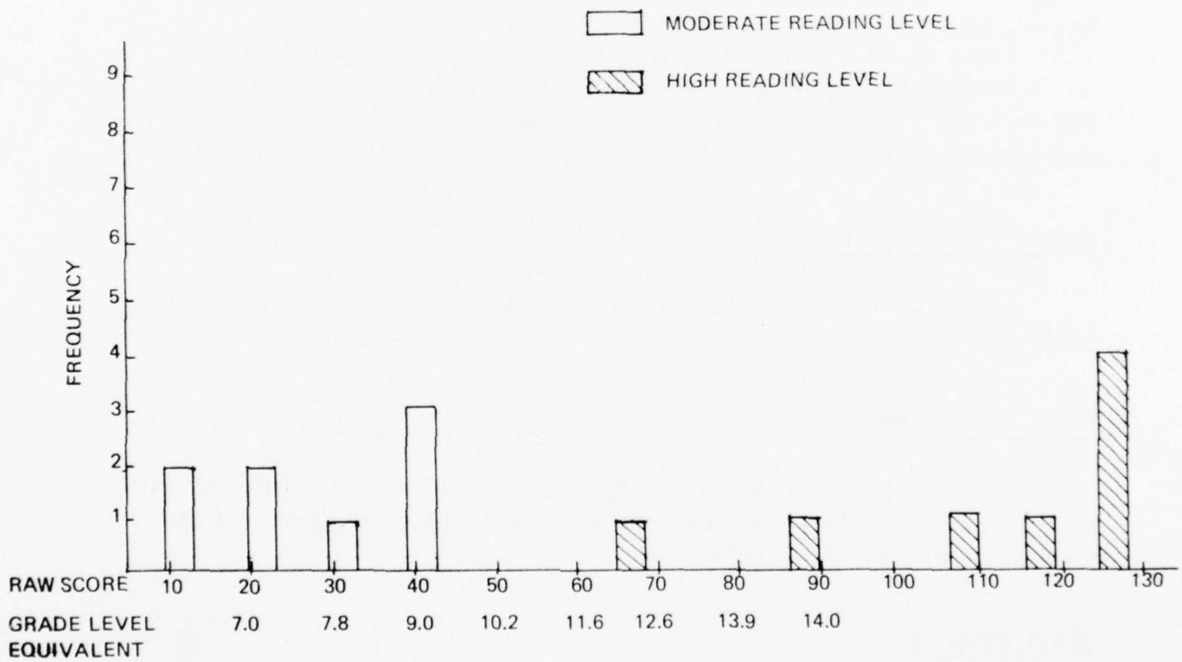


FIGURE 25. DISTRIBUTION OF NELSON DENNY RAW SCORES AND GRADE LEVEL EQUIVALENTS OF MODERATE AND HIGH READING LEVEL SUBJECT GROUPS IN EXPERIMENT III

Table 24

Summary of Analysis of Variance, Experiment III

	SS	*df*	*MS*	*F*
LB	.1962	2	.0981	10.780*
MD	.0573	2	.0286	2.531
YD	.0509	2	.0254	1.954
GRADE LEVEL	8.1290	1	8.1290	58.948**
SUBJECT:GRADE LEVEL	1.9308	14	.1379	14.989**
LB x MD	.3824	4	.0956	8.852**
LB x YD	.5562	4	.1390	11.301**
LB x GRADE LEVEL	.1757	2	.0878	9.648**
LB x SUBJ W/IN GRADE LEVEL	.2561	28	.0091	< 1
MD x YD	.0737	4	.0184	2.000
MD x GRADE LEVEL	.0205	2	.0102	< 1
MD x SUBJ W/IN GRADE LEVEL	.3162	28	.0113	1.228
YD x GRADE LEVEL	.0482	2	.0120	< 1
YD x SUBJ W/IN GRADE LEVEL	.3633	28	.0130	1.413
LB x MD x YD	.9611	8	.1201	13.054**
LB x MD x GRADE LEVEL	.2453	4	.0613	5.676**
LB x MD x SUBJ W/IN GRADE LEVEL	.6064	56	.0108	1.174
LB x YD x GRADE LEVEL	.1930	4	.0482	3.919**
LB x YD x SUBJ W/IN GRADE LEVEL	.6871	56	.0123	1.337
MD x YD x GRADE LEVEL	.0131	4	.0033	< 1
MD x YD x SUBJ W/IN GRADE LEVEL	.5648	56	.0101	1.098
LB x MD x YD x GRADE LEVEL	.1240	8	.016	1.739
(RESIDUAL) LB x MD x YD x SUBJ W/IN GRADE LEVEL	1.0258	112	.0092	

* $p < .05$

** $p < .01$

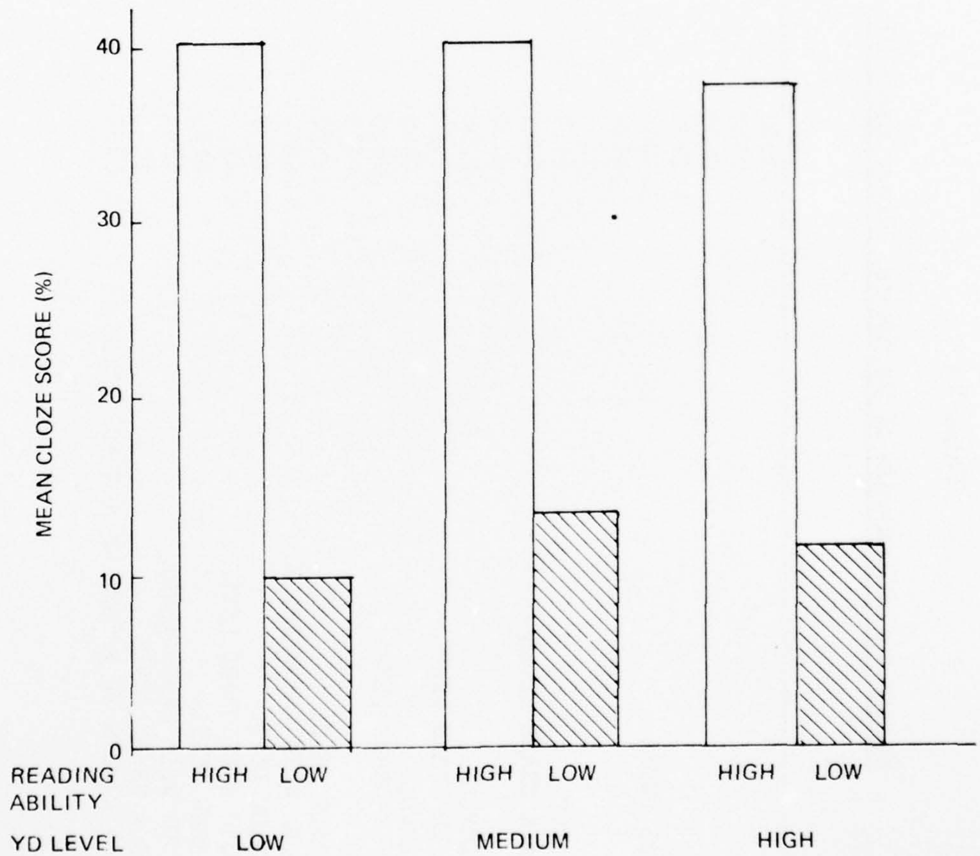


FIGURE 26. MEAN CLOZE SCORE FOR HIGH AND LOW RGL SUBJECTS AT THREE LEVELS OF YNGVE DEPTH

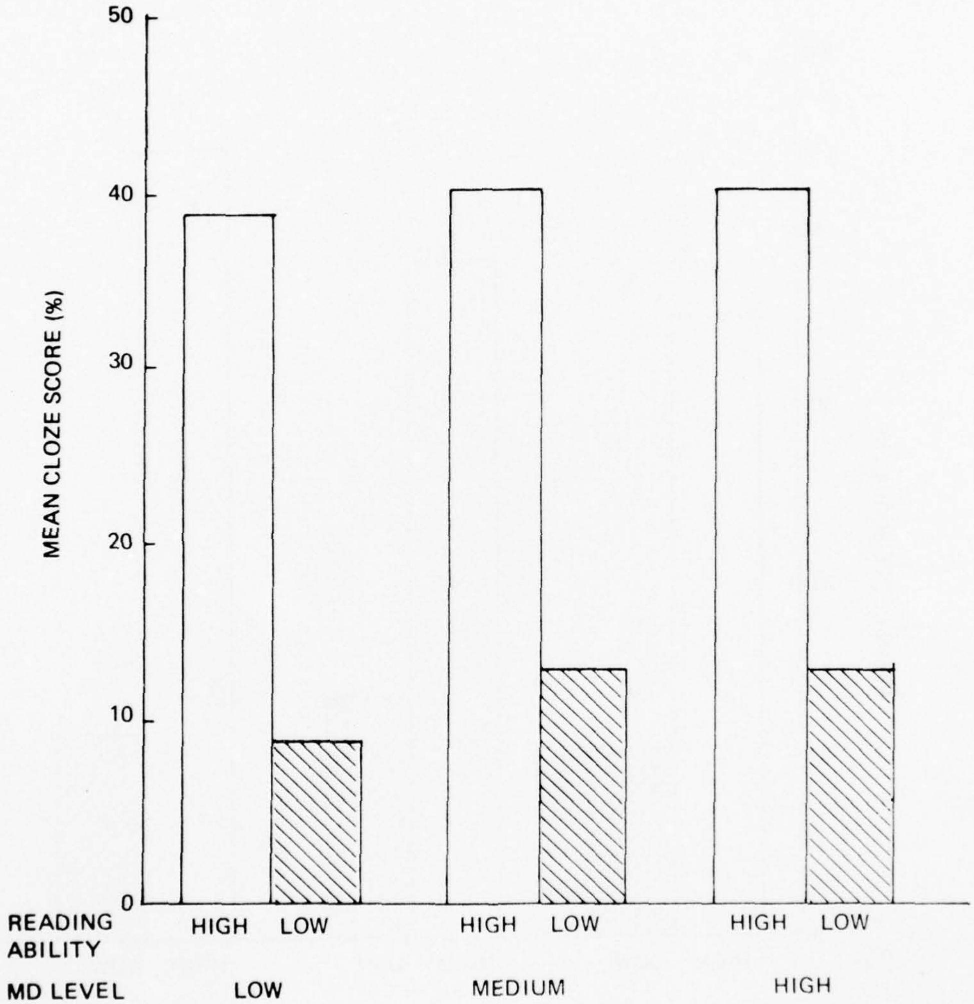


FIGURE 27. MEAN CLOZE SCORE FOR HIGH AND LOW RGL SUBJECTS AT THREE LEVELS OF MORPHEME DEPTH

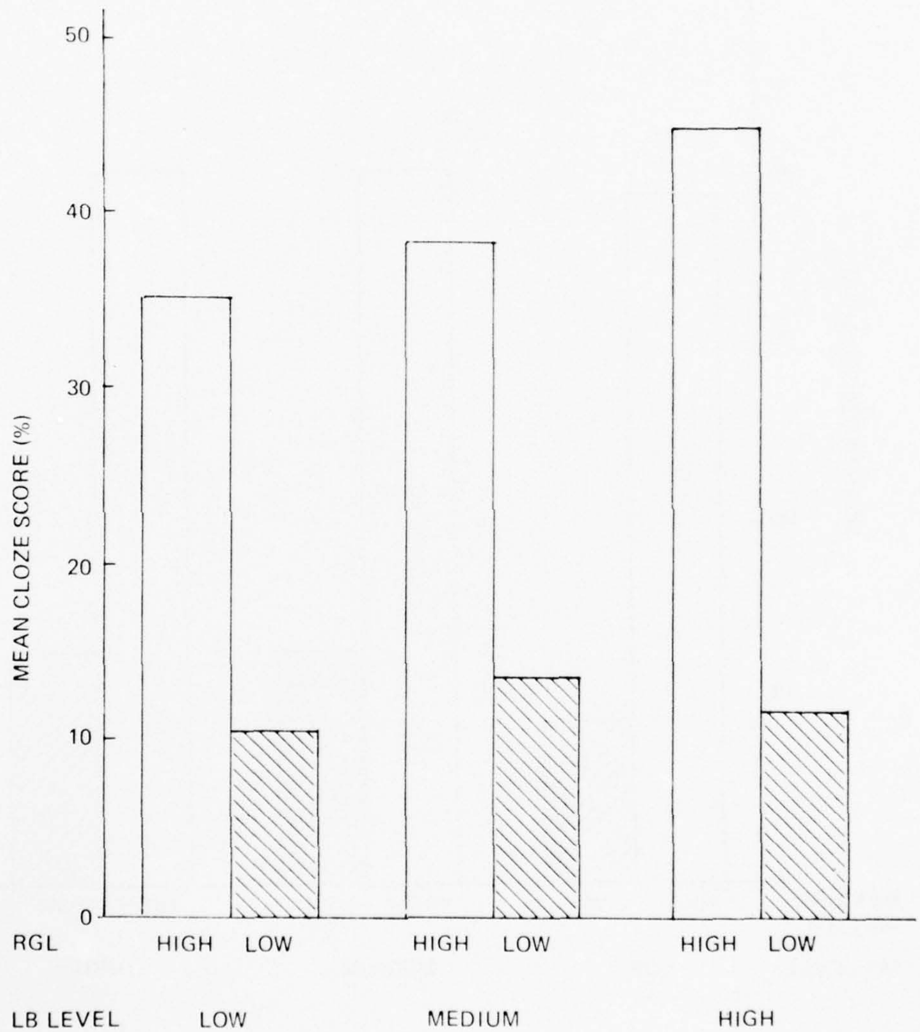


FIGURE 28. MEAN CLOZE SCORE FOR HIGH AND LOW RGL SUBJECTS AT THREE LEVELS OF LEFT BRANCHING

Interactions

As for Experiment II, only two-way interactions between comprehensibility variables are discussed.

In the statistically significant interaction of YD with LB (Figure 29), the effect of variation in LB with YD held at a high level was quite different from the same variation with YD held at low or moderate level. It should be noted that these measures are related in that LB level may be considered to be a contributor to YD; i.e., elevation or depression of LB will tend to move the YD value in the same direction.

The lowest mean cloze score at the low level of LB was that of high YD. Elsewhere, cloze scores associated with high YD were considerably above scores at low and medium YD. The reversal at low LB may be due to the relatedness of the variables, i.e., high YD and low LB are relative opposites. Producing this combination of values required that the writing be somewhat unusual in style. This characteristic may have negatively affected comprehensibility.

The MD by LB interaction was also statistically significant at the .01 level of confidence. This result was clearly due to the difference in effect on comprehension as MD was varied with LB held at the medium level, compared to the low LB and high LB curves. This interaction is presented as Figure 30.

In the MD subexperiment of Experiment I, cloze peaked as a function of MD at the medium value of MD. The same effect was apparent when LB was held at low or high levels. MD may possess the originally anticipated effect only when other variables are held at the medium level.

The interaction between YD and MD was not statistically significant.

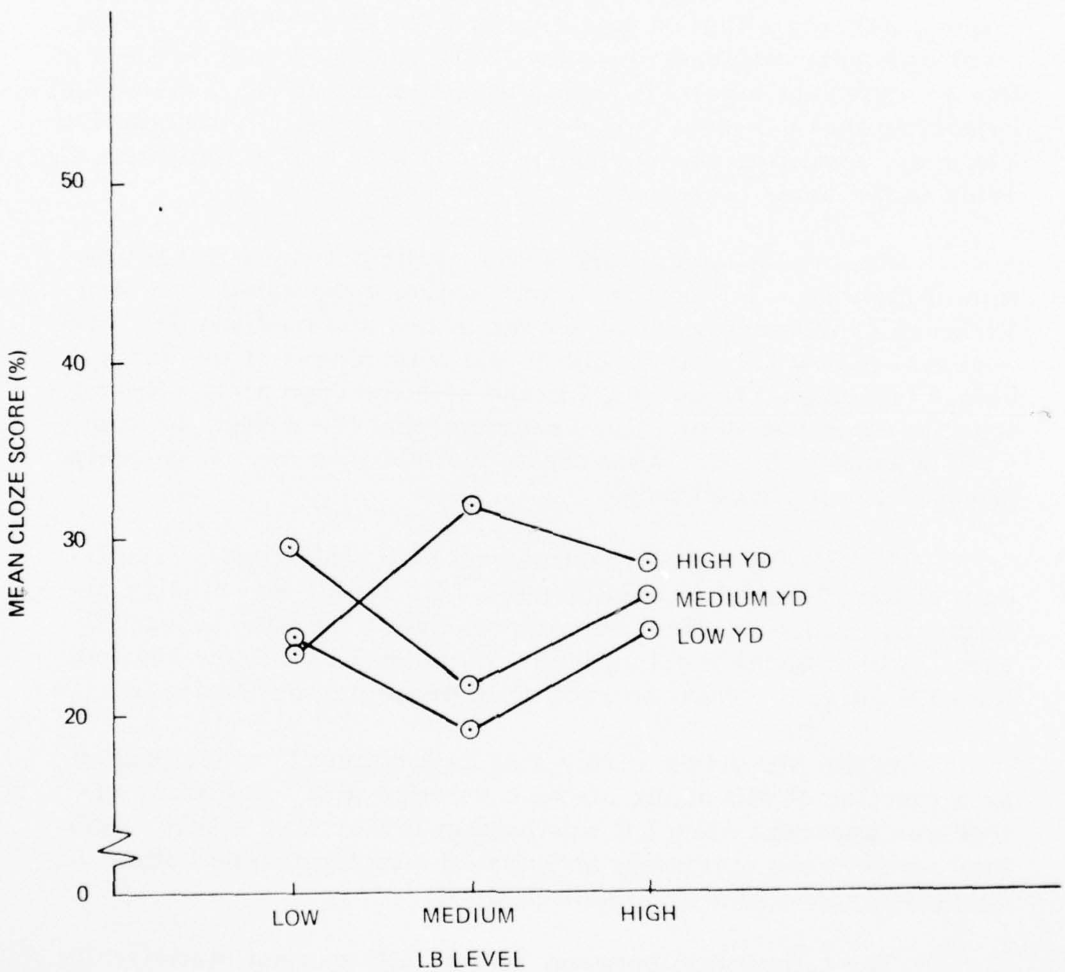


FIGURE 29. INTERACTION OF YD AND LB

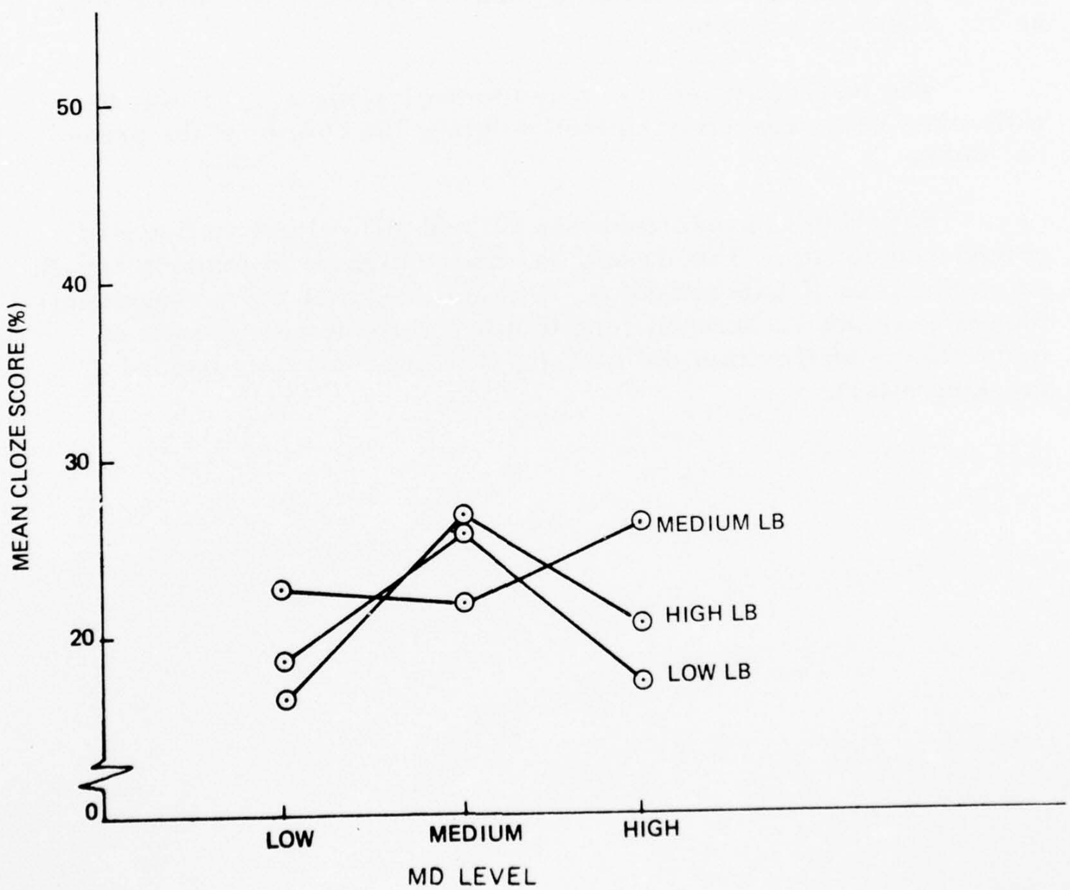


FIGURE 30. INTERACTION OF MD AND LB

Discussion

Statistically significant two-way interactions were evidenced for the psycholinguistically oriented variables, as for the SI oriented comprehensibility variables of Experiment II. As for Experiment I, these findings suggested that a simple additive conceptualization of these measures in combination was not tenable. Such findings are not unique to comprehensibility measurement. Fields such as personality and individual ability differences must also contend with such interactive effects. Such effects suggest that the world of textual comprehensibility measurement is not as simple or orderly as one would like it to be.

The interactive effects noted seem logical and consistent with other data separately collected during the course of the present work.

Experiment I indicated that YD and MD reliably influenced comprehensibility. This result was not confirmed in Experiment III. As in the case of Experiment II, it is believed that the passages employed were not sufficiently long to allow reliable assessment of main effects and/or that the RGL level of the materials caused a masking effect.

Experiment IV

A fourth experiment was performed in order to further investigate one of the variables which had not demonstrated a consistent effect across experiments. Since passage length was posited as a possible explanatory basis for the inconsistency, longer passage lengths were employed in Experiment IV than in Experiments II and III, and several other aspects were varied. CMU was selected for consideration in Experiment IV. The CMU variable had demonstrated a statistically significant effect in Experiment I but not in Experiment II.

The CMU factor is defined as 1 - [NDW(B)/TNW(B)], where NDW(B) represents the number of different words appearing in a textual block and TNW(B) represents the total number of words in the block. The CMU measure is based on examination of text blocks of 100 words. For longer passages, successive blocks of 100 words are measured, and the individual values are averaged.

Subjects

The subjects were eight college undergraduates enrolled in an introductory psychology course. The Nelson-Denny Reading Test (Brown, 1960) was employed to classify the students into two reading ability levels. One group of four students had an average score at the 87th percentile for grade 10 and the other group of four had an average score at the 86th percentile for grade 16. Accordingly, one group of subjects was viewed as high school level readers, and the other group was viewed as college level readers.

Reading Materials

The basic stimuli consisted of four passages of popular literature. The lengths ranged between 458 and 616 words. For convenience, they were designated A_2, A_3, B_2, and B_3.* Paragraph A, a story about cold cures taken from Readers Digest, was written at two extreme levels of word redundancy (A_2, A_3). Paragraph B, a story about the early life of Will Rogers, also taken from Readers Digest, was prepared to reflect the same two extreme levels of type/token ratio (B_2, B_3). Collectively, the passages define two experimental conditions as high complexity (A_3, B_3) and low complexity (A_2, B_2). In order to control the possible confounding of the difficulty of the passages with the CMU measure, all passages were brought to a common RGL (6.5), as measured by the Automated Readability Index (Smith & Senter, 1966). A description of the reading materials is presented as Table 25.

Table 25

Description of the Passages Employed as Experimental Stimuli in Experiment IV

Version of Passage	CMU
Will Rogers	
Low Complexity (A_2)	.52
High Complexity (A_3)	.05
Cold Cures	
Low Complexity (B_2)	.51
High Complexity (B_3)	.04

*The subscript 1 was used to identify the original passage.

Cloze Form Preparation

Cloze forms were developed from the rewritten textual passages in the same manner as for Experiments I, II, and III.

Design and Procedure

The study may be viewed as a two factor mixed design with repeated measures over two paragraph complexity levels, and independent subjects groups to represent two reading ability levels. All subjects were tested in a college classroom during a regularly scheduled introductory psychology class by their regular instructor. During the first portion of the period, each subject was given the Nelson-Denny Reading Test (Brown, 1960) to estimate RGL. Second, each student was exposed to one of the passages with instructions to fill in every deleted word. While no time limit was given, the subjects were told that both speed and accuracy would be scored. The procedure for the cloze data collection for the second passage was the same as for the first passage, and second passage completion followed the first immediately in time.

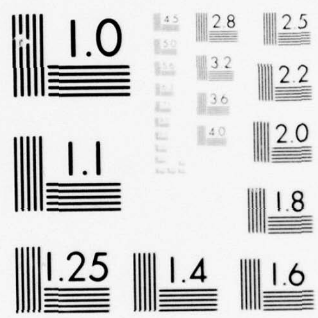
Within the low reading ability group, one of the subjects was exposed to the low complexity version of passage A and then to the high complexity version of passage B. Another one of the subjects at the low reading level received the high complexity version of passage A first and, afterwards, the low complexity version of passage B. The remaining two subjects at the low reading level had a counterbalanced version of the above. Accordingly, the passage sequences were presented to separate individuals of both high and low reading ability to achieve a completely counterbalanced design. Table 26 describes the design of the experiment.

Scoring

Cloze scores were derived employing the same criteria as for Experiments I, II, and III. Completion time data were also collected.

AD-A041 285 APPLIED PSYCHOLOGICAL SERVICES INC WAYNE PA F/G 5/7
STUDIES AND DESIGN SPECIFICATIONS FOR COMPUTERIZED MEASUREMENT --ETC(U)
OCT 76 A I SIEGEL, A R WILLIAMS, W J LAPINSKY F41609-75-C-0037
UNCLASSIFIED AFHRL-TR-76-77 NL

2 OF 3
AD
A041285



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Table 26

Design for Experiment IV

	S	Passage Sequence
Low Reading Ability Subjects	1	$A_2 \longrightarrow B_3$
	2	$B_3 \longrightarrow A_2$
	3	$A_3 \longrightarrow B_2$
	4	$B_2 \longrightarrow A_3$
High Reading Ability Subjects	5	$B_3 \longrightarrow A_2$
	6	$A_3 \longrightarrow B_2$
	7	$B_2 \longrightarrow A_3$
	8	$A_2 \longrightarrow B_3$

Results

Two analyses involving the separate dependent variables were completed. The first was based on the percentage of correctly identified deleted words. The second means analysis was based on the total time taken by the subjects to complete the task. Tables 27 and 28 represent the summaries of these analyses. Graphic presentations of the results appear as Figures 31 and 32. In Experiment IV, statistically significant differences were noted between cloze scores for the high and the low CMU materials.

Table 27

Summary of Cloze Score Analysis of Variance for Experiment IV

Source	SS	df	MS	F
Between Groups	1470	7		
A (Reading Level)	138	1	138.0	n. s.
Subjects Within Groups	1332	6	222.0	
Within Subjects	3791	7		
B (Passage Level)	3452	1	3452.0	55.32**
AB	27	1	27.0	n. s.
B x Subjects Within Groups	312	5	62.4	
Total	5261	15		

** $p < .01$

Table 28

Summary of Working Time Analysis of Variance for Experiment IV

Source	SS	df	MS	F
Between Groups	55	7		
A (Reading Level)	30	1	30	7.1*
Subjects Within Groups	25	6	4.2	
Within Subjects	54	7		
B (Passage Level)	6	1	6	n. s.
AB	2	1	2	n. s.
B x Subjects Within Groups	46	5	9.2	
Total	109	15		

* $p < .05$

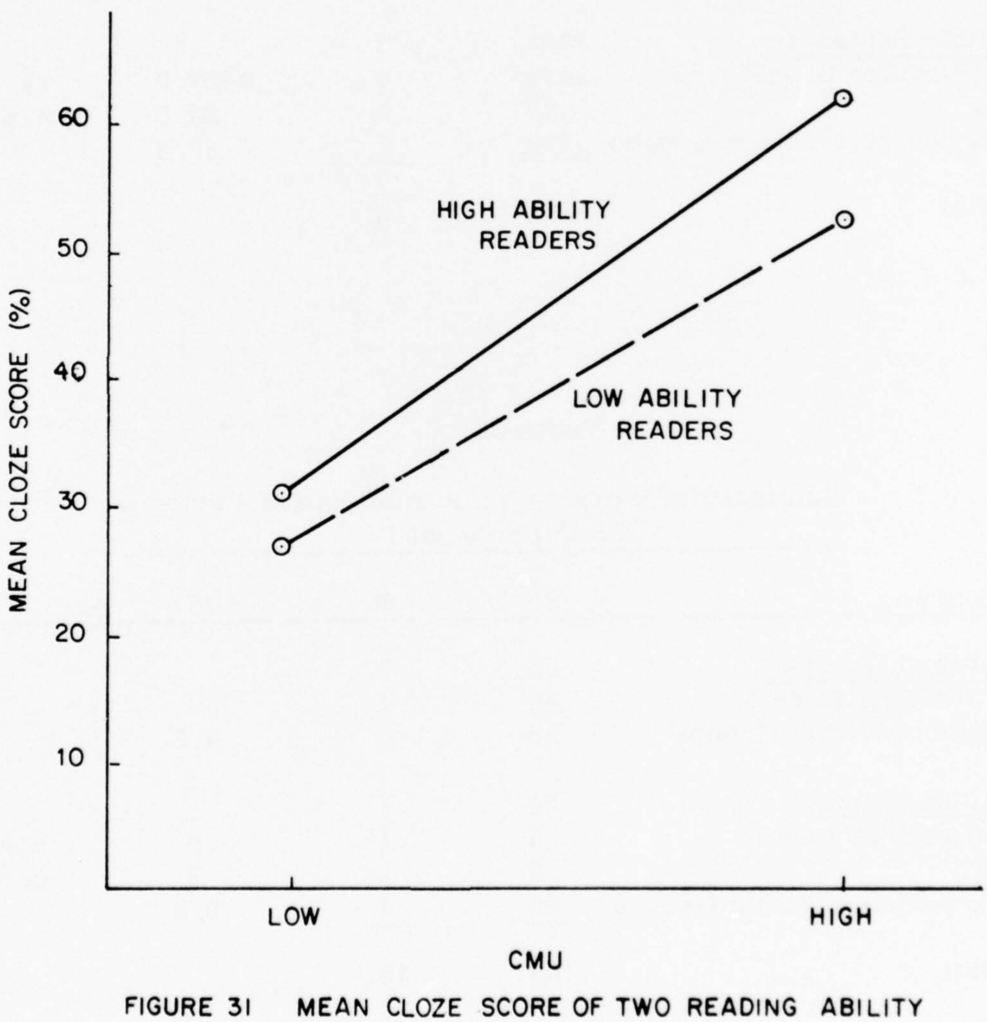


FIGURE 31 MEAN CLOZE SCORE OF TWO READING ABILITY GROUPS AS A FUNCTION OF CMU.

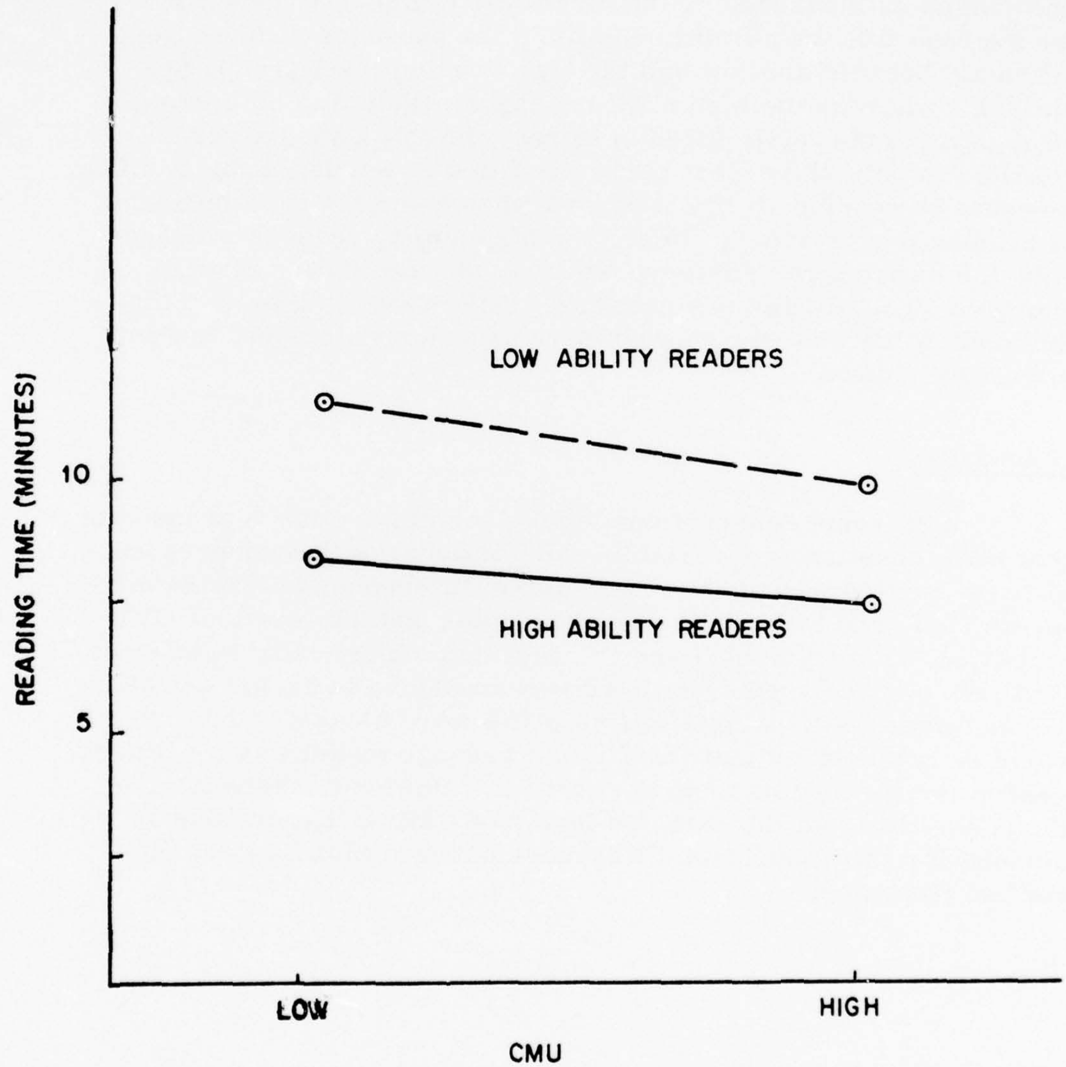


FIGURE 32. WORKING TIME REQUIRED BY TWO READING ABILITY GROUPS AS A FUNCTION OF CMU.

RGL of the subjects did not influence their reading comprehension as measured by the cloze test (Table 27); however, the average time required to complete the passages differed significantly between the low and the high reading level groups (Table 28). Whereas the high CMU passages resulted in an average of 58 percent correctly filled in words, the low CMU passages resulted in only 28 percent correctly filled in words. Subject differences in reading ability appeared when using the time measure of reading performance. Here, the high ability readers averaged only 7.8 minutes per passage, whereas the low ability readers averated 10.6 minutes per passage. It is interesting to note that the high ability readers showed their superiority in speed but not in comprehension.

Discussion

With some reservations, these results are taken to indicate that CMU constituted a variable which influences the comprehensibility of written passages. Two salient differences among Experiments I, II, and IV are the passage lengths and the levels of CMU employed. Experiments I and IV, in which statistically significant CMU effects were noted, utilized textual blocks in excess of 300 words, whereas Experiment II used 135 word blocks. Thus, it could be conjectured that insufficient passage length was a primary reason for the results of Experiment II. However, there is also the possibility that the extreme levels of CMU in Experiment IV account for the significant differences between cloze scores on the two passages.

IV. DISCUSSION AND CONCLUSIONS FROM EXPERIMENTAL FINDINGS

The purpose of the experimental portion of the present study was to verify and clarify the effects of a set of SI and of a set of psycholinguistically oriented variables on the comprehensibility of Air Force technical training and related materials.

Four experiments were performed relative to the first goal. The results of the first experiment were in substantial agreement with prior work, performed under Air Force sponsorship, which indicated that variation of both the SI and psycholinguistically oriented variables does, in fact, affect the comprehensibility of text. The second and third experiments indicated that the SI oriented variables exerted an interactive effect on one another, as did the psycholinguistically oriented measures. Accordingly, the results of these two experiments suggest that comprehensibility cannot be considered to be a simple, additive cognitive attribute. The fourth experiment sought to verify the reason that certain main effects, previously found to exert statistically significant effects on textual comprehensibility, did not produce the anticipated effects in the second/third experiments. With some reservations, due to the nature of the experimental materials used, the results of the fourth experiment might be viewed as supporting the contention that the length of the textual materials employed in the second/third experiments was insufficient. Accordingly, when the present set of results is viewed in association with prior studies (Siegel & Burkett, 1974) investigating the same variables, there is a growing body of evidence supporting the potential of the SI and of the psycholinguistically oriented variables as measures of textual comprehensibility. In this regard, we also point to the high multiple correlations developed by Williams, Siegel, Burkett, and Groff (in press) relative to the power of the present set of variables as predictions of cloze scores (overall $r = .60$; high reading ability group $r = .73$; low reading ability group $r = .46$).

While the mechanism through which these and related variables affect comprehensibility remains to be posited, this mechanism was conjectured at the outset (Siegel & Burkett, 1974) as an

intervening variable called mental load. A relationship was conjectured between the level of these variables in a text and the intellective load that the text places on the reader. Such a relationship would help to explain why a calculus text, although written in small words, might be difficult for a reader or why an electronics manual, although containing somewhat large words, might be highly comprehensible to an electronics technician. Such an intervening variable might also help to explain certain of the interactive effects noted. If, for example, differential intellective loads are imposed by the variables, we would not anticipate a direct, linear stimulus-response relationship between the variables in combination and comprehensibility. Moreover, context may differentially affect the variables. For example, a direct relationship may exist between CMU in an electronics manual and intellective load when the reader is a layman and an inverse relationship may exist for an experienced electronic technician.

The present measures possess certain other advantages. They were based on constructs which are believed to be meaningful and as such, they gain added support. Additionally, they are amenable to objective derivation through formalized counting procedures. Such objectivity removes user bias and judgment from the evaluation of the comprehensibility of a text. The counting procedures make it possible to implement the measures through digital computer techniques.

Thus, the intellective load construct seems to continue to provide a necessary unifying construct for textual comprehensibility analysis. Little was found in the present set of results to negate confidence in this construct. Accordingly, we continue to support the construct for textual comprehensibility measurement purposes.

The interactive effects noted within Experiments II and III are customary in this type of work. While a simple, additive set of comprehensibility variables would represent a desirable goal, such a goal is probably not realistic. It seems more rational to support the use of current variables and supportable measures (as discussed in this report) while the search for new and better variables continues.

We note that the present measures are not normed to RGL. The RGL construct has not proven itself, in our opinion, to be entirely fruitful or meaningful. For example, it is not likely that an intelligent adult reading at the tenth RGL derives the same set of perceptions from a given text as a ten-year-old reading at the same level. Rather, we have normed our measures to other reading materials. It seems more reasonable to compare a given text to a percentile value for other similar texts than to RGL.

The present variables say nothing about format or media. These are considered to be more related to the utility of certain types of written materials (e. g. , operational manuals) than to comprehensibility. Alternatively, format problems which are solved by typographers, table setters, headings, and the like can also be held to lighten the intellective load on the reader and accordingly aid comprehension.

One of the needs of the technical writer is for an objective evaluative device which will tell him not only that a text is at a given level of comprehensibility but also what he can do to increase the comprehensibility of the text. The present method is believed to possess such diagnostic value; however, there has been no evaluation, to date, of the utility of the method in this regard. In a similar vein, technical course directors and managers currently attempt to maintain quality controls and standards for training and other written materials. The CM program, in conjunction with the associated norms, can be employed to establish and maintain minimally acceptable and desirable comprehensibility standards.

With respect to the question of whether or not the present measures are applicative to paper-and-pencil testing materials, little is known. The present set of measures is dependent on textual block sizes of about 300 to 500 words. Available readability measures depend on an adequate sample of text. Yet, some seem (on the surface) to be less dependent on block size than others. Accordingly, there is reason to believe that some selected subset of the present set of measures may be employable for paper-and-pencil test comprehensibility evaluative purposes.

On the basis of the indications of the data, the following conclusions seem tenable:

1. The various Structure-of-Intellect and the psycholinguistically oriented measures, described and investigated, possess potential for measuring and providing a basis for evaluating the comprehensibility of textual materials.

2. The various measures are interactive in nature and the full extent of these interactions remains unknown.

V. THE COMPREHENSIBILITY MEASURES (CM) PROGRAM

Chapter V describes the structure and major logic features for a comprehensibility measures (CM) computer program which could be developed for performing the various textual comprehensibility measurements discussed throughout this report.

The global flow-sequence chart for the CM program is presented as Figure 33. At the outset, we note that the availability of at least one dictionary for use by the CM program is assumed. Also assumed is a version of the text to be processed in machine readable form. These are shown in Figure 33 interfacing with the calculation module since, for the normal run, it is expected they would be assessed by the computer via its bulk storage.

For input, the CM program will accept run specifications in a very flexible author-compatible syntax. A variety of parameters may be entered to determine such items as the length of the text to be measured, the nature and scope of the run, and the quantity and form of output. These may be entered using a common syntax either via interactive terminal (local or remote) or for batch processing (local input or remote batch terminal) depending on the equipment configuration of the computer system for which the CM program may be adapted.

Using these input data specifications and the dictionary and text to be processed, the CM program will perform a run whose purpose is either a dictionary check of selected text (a CHECK run) or the calculation of comprehensibility and other measures (a MEASURE run). Regardless of type of run, various program modules which comprise the CM program will be utilized. These modules are shown in Figure 33 grouped into three categories: operating modules (those dealing with input, program initialization and operator interface), semantic modules (those dealing with calculating the various comprehensibility and reading grade level measures), and result modules (those involved with summarizing and displaying results at various levels of detail). The modules in Figure 33 represent a summary identification; the actual name and functions of each of the 20 program modules is given in Table 29.

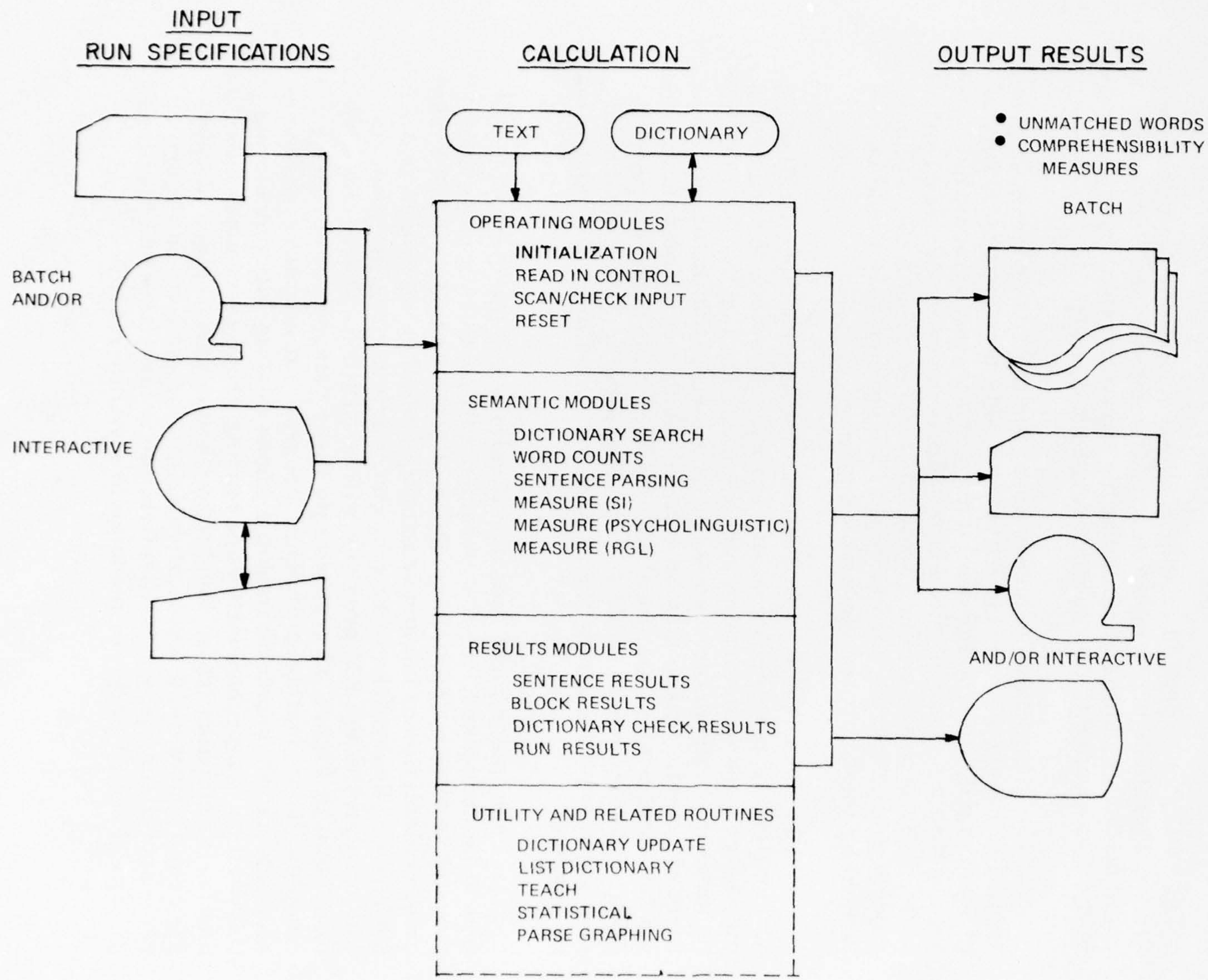


FIGURE 33. OVERVIEW OF THE COMPREHENSIBILITY MEASURES PROGRAM

Table 29

Operational Program Modules in Comprehensibility Measures Program

MODULE IDENTIFICATION	MODULE FUNCTION(S)
1 INITIAL	Perform initializations required for start of run
2 READ	Control read and check of run request input data. Set up run
3 SCAN	Read and scan text for block
4 SCANINPUT	Directly access user input token/record
5 ERROR	Report syntax error in users input
6 RESET	Perform setup (reset) for processing of sentence
7 SEARCH	Check dictionary for word. Extract information or request input
8 COUNT	Maintain running counts for sentence and block summaries
9 PARSE	Parse sentence and determine no. of possible parses
10 MEASURE/SI	Calculate Structure-of-Intellect measures
11 MEASURE/P	Calculate psycholinguistic measures
12 SENTSUM	Cumulate/summarize results of sentence
13 SENTOUT	List/display sentence results
14 RGL	Calculate mechanical reading grade levels
15 RGLOUT	List/display RGL measures
16 BLOCKSUM	Summarize and normalize block results and maintain over the block
17 CHECKOUT	List/display block results of dictionary check
18 MEASUREOUT	List/display block results of measures and regression calculations
19 RUNSUM	Summarize results over all block for run report
20 RUNOUT	List/display run results

Output results will be printed or displayed in the selected level of detail and at the terminal location specified by the user input. Output includes the comprehensibility measures, words not found in the dictionary, and a variety of summarized statistical and parsing results of processing.

A series of utility and related support routines are also identified in Figure 32. These are not specified in detail in this report as they do not come within the scope of the CM specifications directly. They are mentioned here only as a checklist reminder that their availability would be very helpful in an operational environment in which the CM program is expected to be used. Elaboration on the functions of the utility subroutines is given in Table 30.

Table 30

Utility and Support Program Modules Relating to the Comprehensibility Measures Program

Model Identification	*Module Function(s)*
101 UPDATE/D	Add, delete, modify dictionary contents; calculate statistical summaries
102 LIST/D	List/display dictionary contents in a number of ways
103 TEACH	List/display user instructions on CM program runs
104 CMSTAT	Maintain statistics on CM runs (time, words, measure sums)
105 GRAPHPARSE	Print a graphical representation of the structure of selected sentences on the line printer

Also associated with the use of the CM program, but not part of it are programs likely to be already available in a computer facility which supports text editing, word processing, and similar functions. Examples of these programs and their functions are:

> Text entry--generate text files in the format required by CM for comprehensibility measurement
>
> Text editing or update--modify and/or correct a file containing the text to be measured
>
> Test listing/display--record the current revision of a text file for author/editor review and/or publication with or without line numbers

Figure 34 shows the global logic of the CM program in a somewhat more detailed form. Each box, or group of boxes in the chart, presents the program module name, number, and function, and the interrelationships among modules. Prior to discussing this chart, the concepts of a block of text and of a computer run are elaborated.

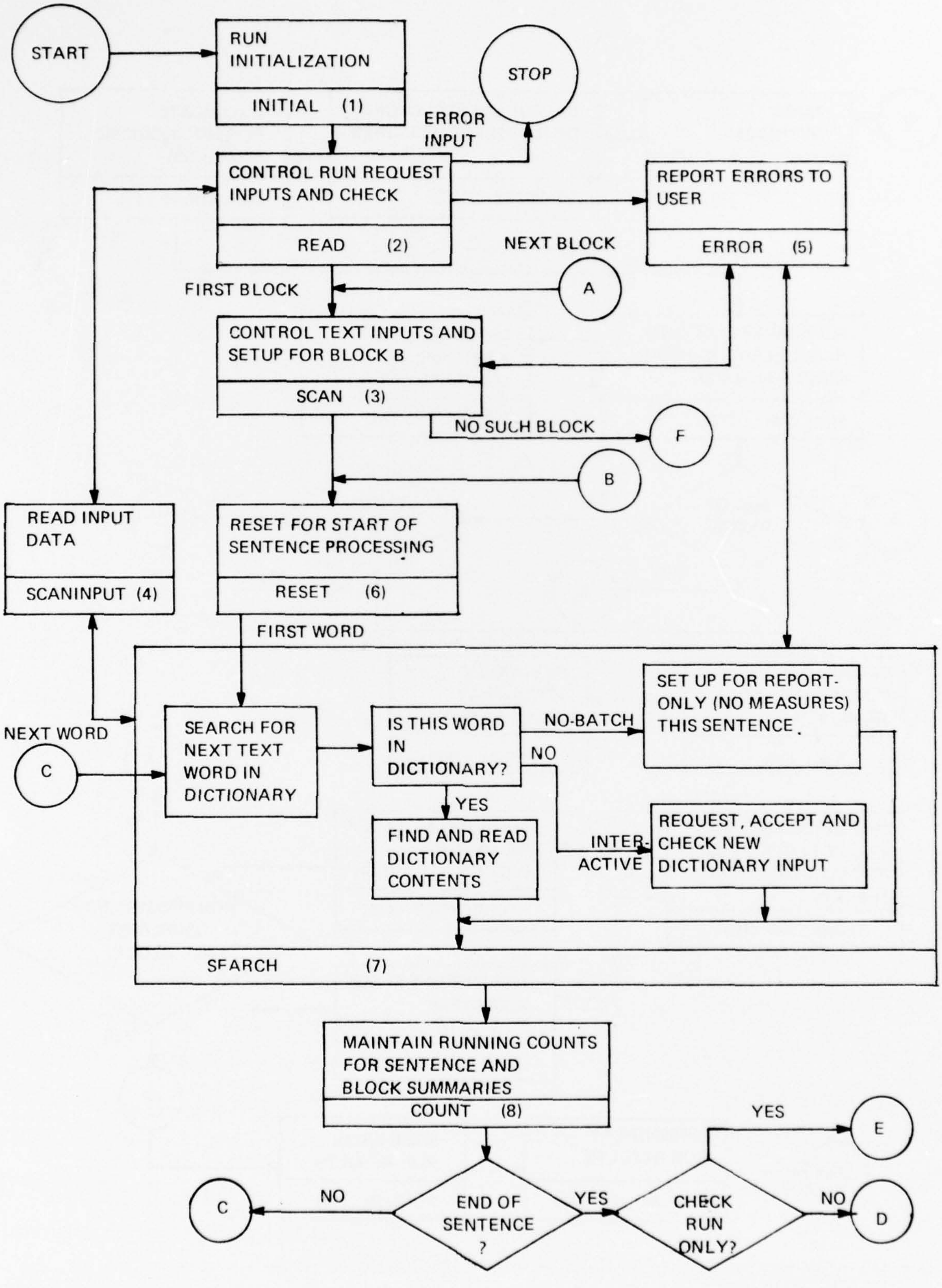


FIGURE 34. COMPREHENSIBILITY MODEL GLOBAL FLOW LOGIC (Page 1 of 2)

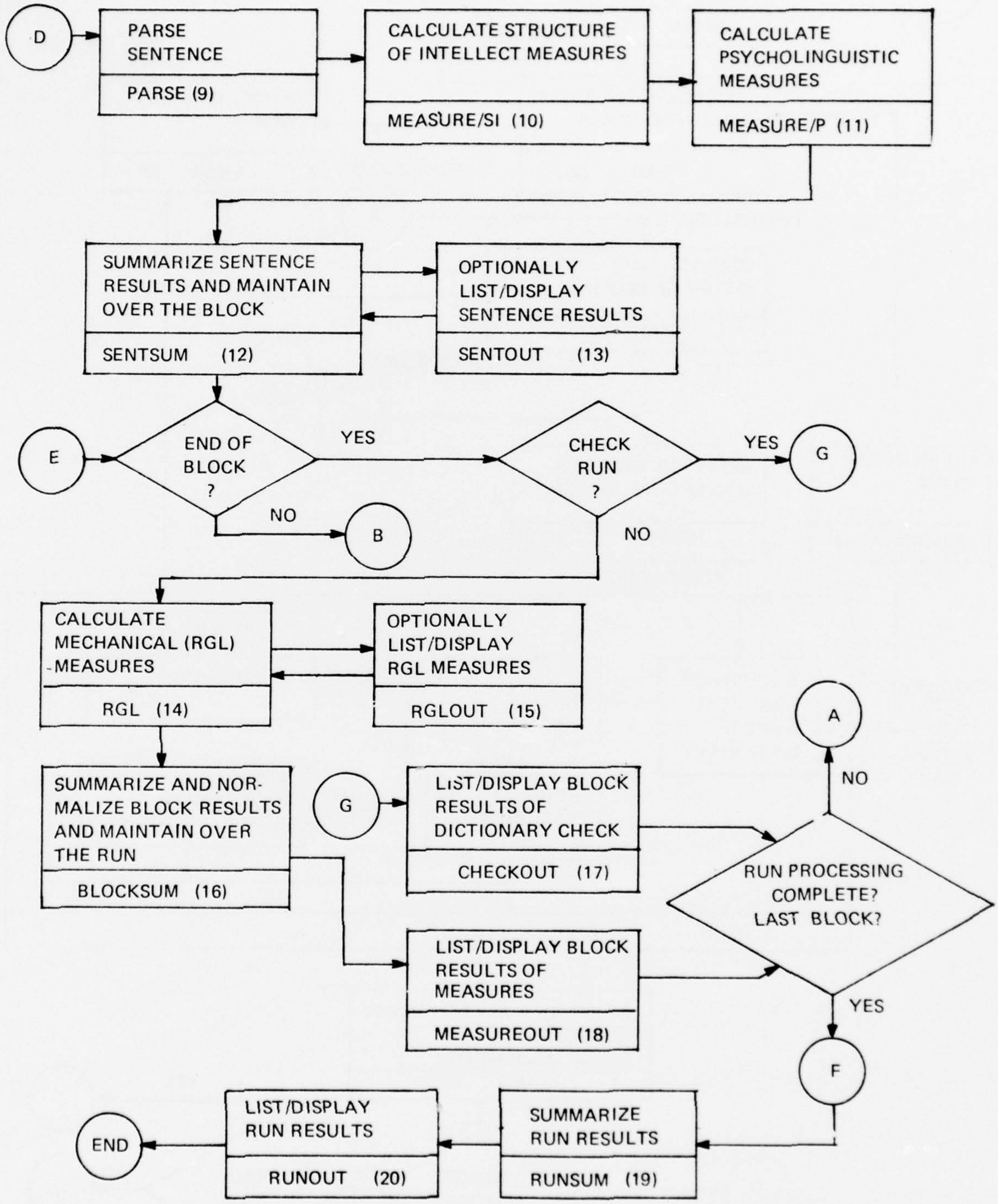


FIGURE 34. COMPREHENSIBILITY MODEL GLOBAL FLOW LOGIC (PAGE 2 OF 2)

Text Block

A "block" of text is a contiguous string of words, abbreviations, and numbers comprising sentences, and whose starting and ending points or length are specified by the CM user. A block must start at the beginning of a sentence and end at the end of a sentence. The minimum block size is arbitrarily specified to be 100 words. Accordingly, the user may specify text blocks over 100 words. However, since a block may not end in the middle of a sentence, the CM program (not the user) will determine the actual number of words, per block, and this will vary from block to block. For example, if the user specifies 500 word blocks, they will be near 500, but not necessarily equal to 500. The user, alternately, has the option of inserting "block mark" codes into his text. This will allow him to calculate comprehensibility measures by page of text, by paragraph, chapter, or the like. However, placement of a block mark in the middle of a sentence is not permitted.

Computer Run

A computer "run" is defined as a series of iterations through the CM program. Each iteration calculates the comprehensibility measures, etc., for a block until the entire specified text has been so processed and an end-of-run summary of all blocks has been calculated and either listed or displayed.

Equipment and Systems Software Requirements

The CM program was planned and designed to be implemented on the CDC Cyber 73-16 computer system at the Technical Training Division of the Air Force Human Resources Laboratory. Languages currently available are PASCAL and FORTRAN. The major equipment features are:

1. 98, 304, 60-bit words central memory
2. 10 peripheral and control processors of 4096, 12-bit words
3. 2 CRT displays with keyboard for operator console
4. extended core storage of 503, 808, 60-bit words
5. disk storage 472 million, 6 bit characters, 30 ms access
6. line printer 1200 lines per minute
7. card reader at 1200 cards per minute
8. 4 magnetic tape units, 9-track, 800 and 1600 bpi, 80, 000 and 16, 000 8-bit characters per second transfer rate
9. Plato terminal

However, the CM program design and specifications were developed for any medium to large scale digital computer system which has peripheral capability for at least the following:

1. a CRT terminal, with keyboard or card reader
2. 1 magnetic tape unit
3. 1 high speed line printer
4. intermediate access storage adequate for the text to be measured, the dictionary and the operating system, discussed subsequently

To be compatible with the computing system available at the Air Force Human Resources Laboratory, the CM program should be implemented in the PASCAL programming language.

Storage Requirements

The central memory requirements (these estimates include a 20 percent reserve factor) of the program are estimated to be:

5, 000 words for global data (COMMON), files, and buffers
1, 200 words for global code
9, 600 words maximum size of any one module

Therefore the maximum central memory required at any one time should be 15, 800 words.

The total estimated storage requirement for all modules of the CM program are expected to be about 50, 000 to 60, 000 words. The following assumptions were made in developing the above estimates:

1. There will be 10, 6-bit characters per word.
2. These will be 10 words for each file as file control information.
3. There will be 2 buffers and a record area for each file.
4. The ability to overlay each module will exist so that only one module (or group of modules) need be in core at one time.

In addition, data files (itemized in Appendix E-1) will require storage in mass memory.

Run Requests

For each computer run of the CM program, the user will select and enter a variety of run request information. Runs are of two types: (1) a CHECK run, to search the selected text and check whether or not the text words appear in the specified dictionary file, or (2) a MEASURE run in which the comprehensibility and other measures are determined. The specific syntax for calling each of these types of runs is displayed in detail in Appendix C. A summary of the type of input required for both types of runs is presented in Table 31 for reference purposes. In Table 31, an asterisk indicates those input types for which default values will be supplied by the computer program if not given by the user, and the default column summarizes the condition if no input is given. The only mandatory input is that needed to identify the text input file. The large number of optional inputs facilitates versatile run requests by the user.

Dictionary File Requirements

A dictionary file is one of the major input files used by the CM program. This file will need to be developed as part of the program development effort since no dictionary is known to exist which contains all of the various required data. Table 32 summarizes the principal information relating to the dictionary. Part A of the table shows the eight items of information which are stored for each word in the dictionary. Seven of these are provided for each dictionary entry. The eighth is updated by the CM program itself.

Part B of Table 32 presents header information relating to the specific dictionary in use; these items are maintained by the CM program. It is anticipated that various users of the CM program will find that more than one dictionary will be required in order to accommodate runs for various texts. (Of course, only one dictionary is required per run.) The CM program provides for selecting a single dictionary by name on each run. Specialty dictionaries may result from continued use of the CM program. Such dictionaries will serve to reduce computer run times since, in general, smaller dictionary sizes will result in shorter run times.

Other Files

Part C of Table 32 presents the two additional files required as CM program input. In conjunction with the Run Request Syntax, a file of words which play the role of introducing explanations is required, as well as a file containing cliches which may occur in the text.

Table 31

Run Request Input List

INPUT TYPE	PURPOSE OF INPUT	INPUT UTILIZATION CHECK RUN	INPUT UTILIZATION MEASURE RUN	DEFAULT
TEXT SPEC	identifies text input file	yes	yes	none-must be input
*START PLACE	starting word line no. of text input file	yes	yes	start of file
*END PLACE	ending word line no. or no. of blocks	yes	yes	end of file
*MODE	batch or interactive run	yes	yes	batch
*DICTIONARY SPEC	identifies dictionary to be used	yes	yes	standard dictionary
*LIST LOCATION	specific printer or terminal output	yes	yes	printer
*COMMENT	report title	yes	yes	"Comprehensibility Measures Program"
*ABORT SPEC	run abort threshold (% not in dictionary)	yes	yes	no threshold or abort
*SAMPLE SPEC	rate of dictionary sampling	yes	no	every word
*BLOCK SPEC	words per block or mark designators	no	yes	500 words
*LIST OPTION	with or without line numbers	no	yes	no text or block nos.
*MEASURE OPTIONS	which measure to calculate and which output to generate	no	yes	all comprehensibility measures and summary output
*EXAMPLE SPEC	identifies 'for example' file	no	yes	standard file
*CLICHE SPEC	identifies 'cliche' file	no	yes	standard file
*PARSE LIMIT SPEC	maximum no. of parses per sentence	no	yes	no limit
*NORM SPEC	identifies type of document	no	yes	overall document type
*SUBJECT READER SPEC	high, low or high and low skill reader	no	yes	all three reader types

**This input is optional; if it is not provided by the user, the CM program will assume and utilize some predetermined default value, as described in the default column.*

Table 32

Dictionary Requirements

A. CONTENTS: the following items will be stored for each word, symbol and abbreviation expected to be encountered.

Source	*Item No.*	*Item*	*Data Type*	*Name*	*Range*
	1	Number of morphemes	N	NOMORE	1 to 9
	2	Number of parts of speech	N	NOPARTS	1 to 9
	3	Parts of speech	A	PART(P)	1 to 9
INPUT	4	Number of syllables	N	NOSYLLABLES	1 to 15
DATA	5	Negativity indicator	A	NEGIND	P or N
	6	Symbolic abbreviation indicator	A	SYMBOLIND	Y or N
	7	Number of words	N	NOWORDS	1 to 5
CM Program	8	Number of references to word	N	NOREFS	0 to 9999

B. HEADER INFORMATION: the following data will be maintained by CM for each CHECK or MEASURE run.

Source	*Item No.*	*Item*	*Data Type*	*Name*	*Range*
INPUT	1	File name (file ID)	A	FILEID	15 symbols
INPUT	2	Dictionary specialty type	A	SPECIALTY	15 symbols
CM Program	3	Last change date	N	DICTDATE	XX-XX-XX
UPDATE/D Utility	4	Total no. of parts of speech - all words	N	TOTAL PARTS	XXXXX
CM Program	5	Total number of references made to all entries	N	REFERENCES	XXXXXX

C. FILES: the following files will be maintained in support of the dictionary.

Source	*File No.*	*File Name*	*Name*	*Size of File*
input	1	Explanation introducer words	EXPLANITRO	30 phrases each 30 AN symbols
input	2	Clichés	CLICHE	10 phrases each 30 AN symbols

Data type code:
N= numeric
A= alphanumeric

CM Program Modules

Table 29 presented the title and general function of each of the 20 program modules comprising the CM program. Each of the module specifications, presented in Appendix B, includes the following information:

NAME:	Full and abbreviated name of the program module
NUMBER:	Serial number of the program module (see Table 29)
PURPOSE:	Brief description of the function of the program module and its purpose in the program.
TECHNIQUE:	Semantic organization or other approach to be utilized in the execution or calculation of the program module.
INPUT:	A list of the coded names of all variables required as inputs to the program module. Refer to Appendix E for a complete list of all data items.
FILES ACCESSED:	A list of the names of all files which are accessed by the program module.
GLOBAL DATA:	A list of all the coded names of all data available to all program modules accessed or changed by the module.
OUTPUT:	A list of the coded names of all variables generated by the program module.
MODULES CALLED:	A list of the coded names of all other program modules which are called by the program module.
CALLING MODULES:	A list of the coded names of other program modules which call the program modules.
FLOWCHART:	A summary flow logic (if required) showing sequencing of major call tasks.
COMMENTS:	Notes on the program module such as difficulty, size, importance, or the like.

Although not explicitly stated in Appendix B, it is assumed that the structured programming technique will be utilized in the development of each of the CM program modules. As part of the approach, to facilitate program checking a queue-trail list giving the module numbers in the sequence in which they are called during a run should be introduced, and a count of the number of times each module is entered should be maintained.

A running commentary of the total program may be seen by simultaneously reviewing the Appendix B specifications, principally the purpose and technique sections, while referring to the global flow chart, Figure 33.

Sentence Processing

Each sentence is processed independently. This is done by searching for each word's record in the dictionary file (SEARCH module) and then maintaining the principal sentence-oriented running tallies required for the calculations of the comprehensibility measures (COUNT module). The process repeats for all words in the sentence and when completed, the processing of the total sentence proceeds with sentence parsing (PARSE module), and calculations of measures (Measure/SI, and Measure/P modules).

Parsing

While most of the SI measures may be obtained using numeric processing, several in the psycholinguistic category require actual parsing of each sentence of the text. Based on the state-of-the-art in automated sentence parsing, automatic sentence parsing represents one of the most difficult aspects of CM program implementation. Hays (1967) reviewed the basic techniques of parsing, with an emphasis on implementation. Aho and Ullman (1972) also presented a thorough review of the various parsing techniques, emphasizing the type of parser which is most appropriate for languages of various descriptions. Yet, to the best of our knowledge of the current technology, no automatic parser exists which will meet the requirements of the CM program.

To date, a completely general system which will "understand" arbitrary input text in English and act on this understanding is well beyond technological capability. Even the limited capability to parse large number of general sentences correctly is unavailable. In routine auding and reading tasks a person uses the context in which information is presented to resolve ambiguities, but this process requires such a wealth of specialized information that it is all but impossible to conceive of automating it completely.

Computerized systems have been successfully developed, at least on a pilot scale, which have structures containing some specialized information at their disposal, and as a result can extract unambiguous meaning when related to a small part of a single subject.

However, a general system has not been devised. However useful a parser of the specialized type may be in other applications, parsers which operate only on specialized input texts have serious handicaps for use in the CM program. The main disadvantage for general work as well as for the CM program is the fact that they do not work well with ordinary English. Their versions of the language are often "telegraphic English," in which words are omitted, coded, and/or abbreviated. In other cases, such parsers employ a quasi-logical application-unique notation, and in still other cases, highly restricted dictionaries or grammar rules are involved. These are clearly inapplicable to the CM program, whose goal is to process any English prose text, as written.

A parsing technique was sought which will produce all possible grammatically legitimate parses of ambiguous English sentences. The technique developed will be applied within the CM program to each sentence of each text block to be measured for comprehensibility. The principal output from the parser will be the identification of all sentence elements (parts of speech, clauses, phrases, etc.) for each sentence.

To illustrate the ambiguity problem, consider Figure 35, which attempts to disambiguate the sentence: "Never place your fingers in the cutting area." The second column of Figure 34 presents the potential parts of speech for each word in the sentence. These are obtained in the CM program by the relatively simple expedient of a dictionary lookup, which would be required regardless of the parsing scheme. This sample eight-word sentence has many more than one possible parse because some of its constituent words have multiple parts of speech potential. The third column of Figure 32 indicates that up to 1x2x2x1x1x2x3x1 = 24 parses are possible without bringing sentence grammar logic to bear. The main task of the parser, then, is to reduce the number of potential parses to one or a relatively small number.

As a byproduct, the parser should identify selected parts of speech and generate a tree structure. It must be able to handle context-free phrase structure grammar, where the logical rules of the grammar are built into the parsing program module. The main output will be the selection and identification of the most likely part of speech for each word in the sentence. In the example, the desired parser output would be the underlined part of speech in the second column.

The parser, however, must also produce its output in a form suitable for use as input to other program modules which will accomplish additional calculations. These parsing requirements generated by other modules are identified below:

Sample Sentence Words	*Possible Parts of Speech*	*No. of Possible Parts of Speech*
Never	adverb	1
place	noun, verb	2
your	pronoun, objective	2
fingers	noun	1
in	preposition	1
the	article, adverb	2
cutting	adjective, noun, verb	3
area	noun	1

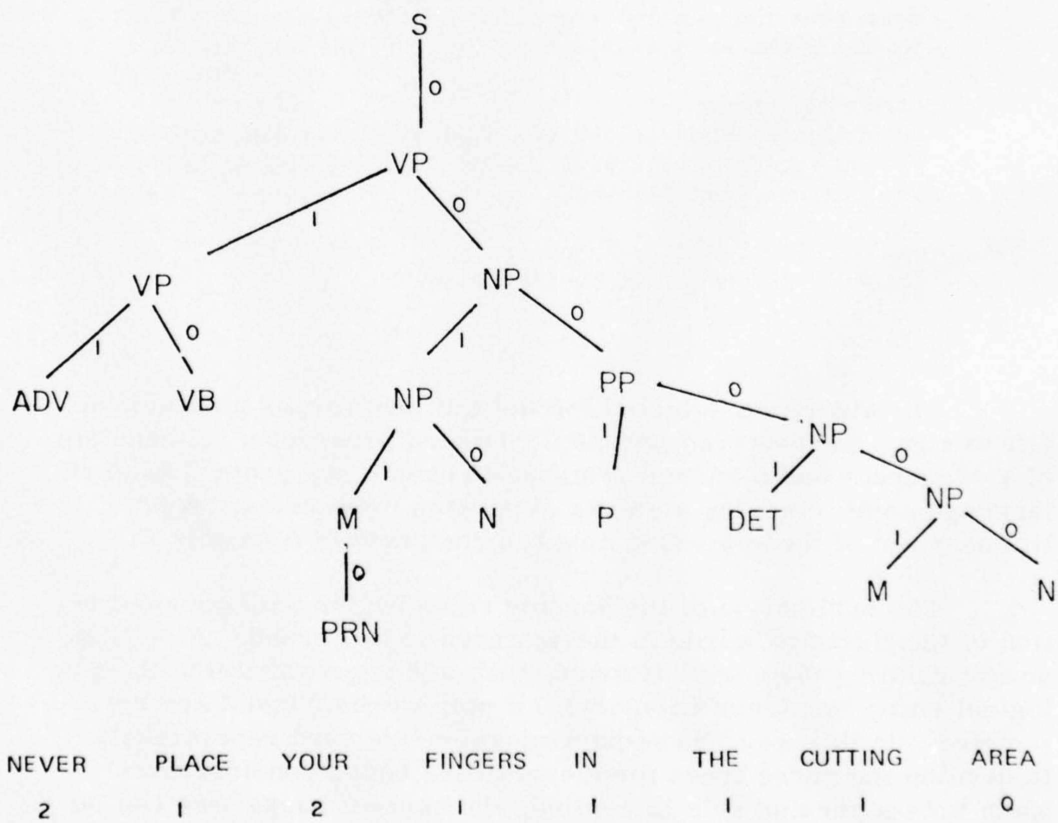


FIGURE 35. POSSIBLE PARSES FOR SAMPLE SENTENCE

YD- Yngve Depth -
Determine the tree structure of the sentence being parsed, and assign 0 or 1 states to the segments as shown in the Figure 34 example. This requirement alone implies that a phrase structured, context free parsing scheme is required.

TC- Transformational Complexity
Determine whether the main clause verb is of a form which causes it to be considered active, passive, active-negative, or passive-negative.

CE- Center Embedding
Determine the number of phrases to the "right" of the subject verb in the sentence.

RB/LB - Right Branching and Left Branching
Determine the number of chained modifying clauses on the right of the object or left of the subject.

DC- Deleted Complement
Determine whether or not the subject of the sentence is the object of a modifying clause in which the relative pronoun has been deleted.

CMR- Cognition of Semantic Relations
Identify the nouns in the sentence.

The algorithm selected for the CM program is a bottom-up, left to right, context free phrase structured procedure. It consists of a preparser and a parser designed to handle any general English language sentence using a series of parsing logic rules. A preliminary set of these parsing rules is contained in Appendix F.

The application of the parsing rules begins with consideration of the first two words in the sentence to be parsed. A rule is sought relating them and, if found, they are aggregated into the new logical entity (sentence category). If not, words 2 and 3 are attempted. In this way, these parsing rules are used repetitiously to develop the parse trees for the sentence under consideration. Each time a parsing rule is applied, the current parse tree can be extended upwards from the words comprising the sentence to the peak made representing the total sentence. Thus, a completed tree can be thought of as a sequence of applications of the parsing rules.

The rules are of the following format:

NP → DET N (a noun phrase may be composed of a determiner and a noun)

which is interpreted as follows: whenever the sequence of categories on the right side of the rule appear in the part of the parse tree which has been completed so far, the tree can be extended upwards from these categories to the category on the left side of the rule. Three data arrays are associated with the set of parsing rules: LEFT, RIGHT, and LENGTH. LEFT contains the categories which appear on the left side of the arrow in the respective rules; RIGHT contains the categories appearing on the right side of the rules, so that the ith member of the array RIGHT contains as many categories as there are categories on the right side of the arrow in the ith rule; and LENGTH, is an array in which the ith member is the number of categories on the right side of the arrow in the ith rule.

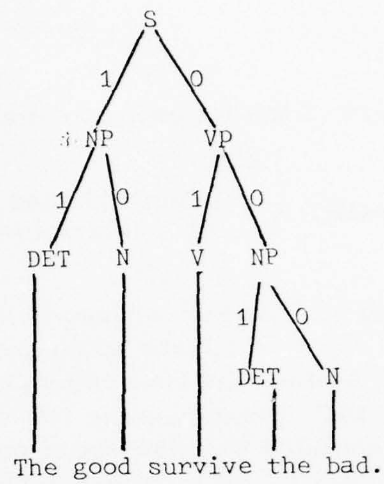
During this process, the program module maintains files for each word of the sentence under consideration, as shown in Figure 36, which illustrates the parsing of the sentence: "The good survive the bad." The parsing module is executed once for each possible combination of assignments of categories to the words of the sentence. The preparser routine partially develops this table on the basis of one of the allowable assignments of categories to words and phrases. The preparser partially generates the columns in the parse table (Figure 36) for category, pointer, signal, and rule, and transfers control to the parser. The appendix contains the specifications for the parsing module.

Sentence Summary

A summary of all sentence results (SENTSUM module) and listing or display (SENTOUT module) concludes the sentence processing. Control reverts to circle B in Figure 33 if the sentence processed is the last one in the text block; this provides for repeating the process for each sentence. If the end of the block has been reached, the traditional measures are determined (RGL module) and optionally recorded (RGLOUT module). Appendix D shows the output formats. The BLOCKSUM module then summarizes the results of each

SAMPLE SENTENCE: The good survive the bad.

TREE



Yngve depth: 2 1 1 1 0

Simplified PARSE-TABLE

Average Yngve Depth= 5/5= 1

	Category	*Pointer*	*Signal*	*Rule*	*YD#s*	*Source*
1	Det	6	1	0	1	Original Sentence
2	N	6	-1	0	0	
3	V	8	1	0	1	
4	Det	7	1	0	1	
5	N	7	-1	0	0	
6	NP	9	1	4	1	Pre-parser
7	NP	8	-1	4	0	
8	VP	9	-1	3	0	Parser
9	S	0	0	1		
	Taken from dictionary	Points to next category above	1= left 0= other	identifies rule used		

RULES:
1= S → NP VP
2= NP → N
3= VP → V NP
4= NP → Det N

ABBREVIATIONS:
NP-noun phrase
N -noun
VP-verb phrase
V -verb
S -sentence
Det-determiner

To calculate average Yngve depth, for each path from the top-most S of the tree to the individual words of the sentence, add the YD#s which have been assigned to the branches of the tree as above. Then add together all these sums and divide into the number of words in the sentence.

Figure 36. Example of parsing technique.

sentence and generates data required to be displayed in the MEASURE-OUT module. The normalization of measures into percentiles over the 5th to 95th percentile range is included in BLOCKSUM. The tables used in this process are based on actual measurement of Air Force technical materials and are given in Appendix G. Using these percentile scores for each measure, BLOCKSUM also calculates the composite index representing a single comprehensibility score for the entire textual block.

Processing then returns to circle A of Figure 33 to perform resets and process the next textual block. When the last block has been processed, the RUNSUM module develops all summary data required to list or display results of the entire run for all blocks. The RUNOUT module controls this display or listing to complete the run processing.

Output Results

Formats for each of the five types of output are specified in Appendix D:

Figure	Types of Output	Automatic/Optional
D-1	Sentence Summary	O
D-2	Dictionary Check	A (CHECK RUN)
D-3	Block Summary	O
D-4	Run Summary	A (MEASURE RUN)
D-5	RGL Summary	O

On these reports, the line and word number are given to orient the user, allowing him to correlate the computer results with any given sentence or block of text by number. When requested, the dictionary check output displays all words in text which were searched in the dictionary and found not to be included. A given word is listed only once in this output, together with the number of times it occurs in the block. The percentage of words not found in the dictionary is listed for each block and for the total run. The space provided at the right of the listing/display is for comment by the analyst or editor. This provides the analyst with a convenient location for entering his instructions for future processing. On a word-by-word basis,

the analyst can enter his decision to make a:

	Code
spelling change in text	1
new dictionary entry	2
dictionary word change	3

This same listing then can be used as an input for required changes in text or dictionary before another run is made.

The block results shows values and percentiles for each measure and for the major variables, together with the composite indices. The text will be printed (with or without line numbers) if so requested in the LIST option of the run request.

The results for the run present all measure values both as calculated and after conversion into percentiles on a block-by-block basis. Complete indices are also shown. Some specific tallies of useful variables (number of words per block, average number of parses per sentence, number of sentences parsed and not parsed, average number of morphemes per word) are also specified. The last report of RGLs and the ARI index is optional but similar in organization to the run summary report.

<u>Error/Condition Messages</u>

A variety of conditions may arise during a CM program run which terminate a run or which represent situations about which the user must be made aware. A list of these conditions and the specific wording for the messages to be recorded or displayed are specified in Table 33. Messages will be displayed on the user's interactive terminal or recorded in the printout, depending on whether the run request syntax <LIST LOCATION> is either TERMINAL or PRINTER (default is printer).

Table 33

Error Message Identification

	ERROR MESSAGE	SOURCE MODULE	ACTION CODE
1	MISSING "CHECK" OR "MEASURE."	READ	A
2	UNIDENTIFIED REQUEST.	READ	C
3	INVALID <FILE ID>.	READ	C
4	NO SUCH TEXT LINE.	SCAN	E
5	NO SUCH WORD NUMBER IN TEXT LINE.	SCAN	E
6	FILE NOT DICTIONARY FILE.	READ	D
7	FILE NOT PRESENT.	READ	D
8	FILE NOT TEXT FILE.	READ	A
9	FILE NOT EXAMPLE FILE.	READ	D
10	FILE NOT CLICHE FILE.	READ	D
11	SYNTAX ERROR IN < > SPEC.	READ	C

ACTION CODE:

A= ABORT RUN
E= Produce final reports and terminate
D= Use default file; abort if default not present
C= Continue syntaxing input, then abort

Wherever possible, the program is forced to continue rather than be terminated at the point the error or condition is recognized and the message given. Table 33 identifies those errors or condition which should result in programmatic termination, and also specifies the conditions under which the CM program will continue.

Discussion of CM Program Development

The present work made considerable progress in specifying the characteristics of a high speed digital computer program which would allow automatic calculation of the various measures. Such a program is detailed in the appendices to this report, and its development seems entirely possible. The required parsing subroutine represents the most difficult aspect of such a program. However, the logic for such a program was developed and seems to be adequate for the purposes on hand. The decision to develop a special parsing program was made on the basis of the advice of linguistics and computer analysts who indicated that known and available parsing programs would: (1) not entirely fill the current requirements, and (2) consume excessive computer running time. Hence, the tailored parsing program, as compared with an "off the shelf" program, was believed to be more cost effective in the long run.

Other aspects of the CM program are relatively straightforward and seem to possess little risk. Hence, completion of the programming aspects of the required work seems possible.

The computer program as designed will measure textual characteristics and provide diagnostic information. It will not suggest alternate wordings or sentence constructions. Such decisions are best left to the technical writer. This point of view has been expressed recently by Sticht and Zapf (1976), who wrote:

> The computer seems to offer potential help, especially if the functions assigned it are those the writer does not like to do, or do well, and that are compatible with computer capabilities. For example, computers can be made to do language recording and storage tasks easily and efficiently, but language decision tasks only with difficulty and poorly, as the machine translation literature clearly shows. Humans, on the other hand, have just the opposite proclivities. Consequently, computer aid to the writer should focus on the former...

REFERENCES

Aho, A.V., & Ullman, J.D. *The theory of parsing, translation, and compiling. Vol. I. Parsing.* Englewood Cliffs: Prentice-Hall, 1972.

Bormuth, J.R. Readability, A new approach. *Reading Research Quarterly*, 1966, *1*, 79-132.

Brown, J.I. *The Nelson-Denny reading test. Vocabulary, comprehension, rate.* (revised) Form A. Boston, Mass.: Houghton-Mifflin, 1960.

Caylor, J.S., Sticht, T.G., Fox, L.C., & Ford, J.P. *Methodologies for determining reading requirements of military occupational specialties.* (Draft Technical Report) Presidio of Monterey, Calif.: Human Resources Research Organization, HumRRO Division No. 3, January 1972.

Coleman, E.B. The comprehensibility of seven grammatical transformations. *Journal of Applied Psychology*, 1964, *48*, 186-190.

Coleman, E.B. Learning of prose written in four grammatical transformations. *Journal of Applied Psychology*, 1965, *49*, 332-342.

Evans, R.V. The effect of transformational simplification on the reading comprehension of selected high school students. *Journal of Reading Behavior*, 1972-1973, *5*, 273-281.

Flesch, R. A new readability yardstick. *Journal of Applied Psychology*, 1948, *32*, 221-233.

Fodor, J. Current approaches to syntax recognition. In D.L. Horton & J.J. Jenkins (Eds.), *Perception of language.* Columbus, Ohio: Charles E. Merrill, 1971.

Fodor, J., & Garrett, M. Some syntactic determinants of sentential complexity. *Perception and Psychophysics*, 1967, *2*, 289-296.

Foss, D.J., & Crains, H.S. Some effects of memory limitation upon sentence comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 541-547.

Gough, P.B. Grammatical transformations and speed of understanding. *Journal of Verbal Learning and Verbal Behavior*, 1965, *4*, 107-111.

Guilford, J.P. *The nature of human intelligence.* New York: McGraw-Hill, 1967.

Guilford, J.P., Comrey, A.L., Green, R.F., & Christensen, P.R. *A factor analytic study of reasoning abilities. I. Hypotheses and description of tests.* Los Angeles: University of Southern California, 1950.

Guilford, J.P., Geiger, R.M., & Christensen, P.R. *A factor-analytic study of planning. I. Hypotheses and description of tests.* Los Angeles: University of Southern California, 1954.

Guilford, J.P., & Hoepfner, R. *Structure-of-intellect factors and their tests.* Los Angeles: University of Southern California, 1966.

Guilford, J.P., & Hoepfner, R. *The analysis of intelligence.* New York: McGraw-Hill, 1971.

Hakes, D.T. Effects of reducing complement constructions on sentence comprehension. *Journal of Verbal Learning and Verbal Behavior,* 1972, *11,* 278-286.

Hamilton, H.W., & Deese, J. Comprehensibility of subject-verb relations in complex sentences. *Journal of Verbal Learning and Verbal Behavior,* 1971, *10*, 163-170.

Hays, D.G. *Introduction to computational linguistics.* New York: American Elsevier, 1967.

Hicks, C.R. *Fundamental concepts in the design of experiments.* New York: Holt, Rinehart, & Winston, 1964.

Hoeptner, R., Guilford, J.P., & Merrifield, P.R. *A factor analysis of the symbolic-evaluation abilities.* Los Angeles: University of Southern California, 1964.

Klare, G.R., Sinaiko, H.W., Stolurow, L.M. *The cloze procedure: A convenient readability test for training materials and translations.* (Science and Technology Division Paper P-660). Arlington, VA.: Institute for Defense Analyses, 1971.

Lambert, J.V., & Siegel, A.I. Psycholinguistic determinants of readability. In A.I. Siegel, & J.R. Burkett (Eds.), *Application of the structure-of-intellect and psycholinguistic concepts to reading comprehensibility measurement.* (AFHRL-TR-74-49). Lowry AFB, Colo.: Air Force Human Resources Laboratory, September 1974.

Martin, E., & Roberts, K.H. Grammatical factors in sentence retention. *Journal of Verbal Learning and Verbal Behavior,* 1966, *5,* 211-218.

Peltz, F.K. The effect upon comprehension of repatterning based on students' writing patterns. *Reading Research Quarterly,* 1973-1974, *9,* 603-621.

Powers, R.D., Sumner, W.A., & Kearl, B.E. A recalculation of four adult readability formulas. *Journal of Educational Psychology,* 1958, *49,* 99-105.

Schwartz, D., Sparkman, J.P., & Deese, J. The process of understanding and judgments of comprehensibility. *Journal of Verbal Learning and Verbal Behavior,* 1970, *9,* 87-93.

Siegel, A.I., & Bergman, B.A. Readability/comprehensibility as related to the structure-of-intellect model. In A.I. Siegel & J.R. Burkett (Eds.), *Application of structure of intellect and psycholinguistic concepts to reading comprehensibility measurement.* (AFHRL-TR-74-49), AD-A001573. Lowry AFB, Colo.: Technical Training Division. Air Force Human Resources Laboratory, September 1974.

Siegel, A.I., & Burkett, J.R. (Eds.), *Application of structure-of-intellect and psycholinguistic concepts to reading comprehensibility measurement.* (AFHRL-TR-74-49) AD-A0001 573. Lowry AFB, Colo.: Technical Training Division, Air Force Human Resources Laboratory, September 1974.

Siegel, A.I., Lautman, M.R., & Burkett, J.R. Reading grade level adjustment and auditory supplementation as techniques for increasing textual comprehensibility. *Journal of Educatioanl Psychology,* 1974, *66,* 895-902.

Slobin, D.I. Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior,* 1966, *5,* 219-277.

Smith, E.A., & Sentner, R.J. *Automated readability index.* (AMRL-TR-66-220). Wright-Patterson AFB, Ohio: Aerospace Medical Division, November 1966.

Sticht, J.G., & Zapf, D.W. *Reading and readability research in the armed services.* HumRRO FR-WD-CA-76-4. Monterey: Human Resources Research Organization, Western Division, September 1976.

Taylor, W. Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 1953, *30*, 415-433.

Wang, M.D. The role of syntactic complexity as a determiner of comprehensibility. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 398-404.

Wason, P.C. The processing of positive and negative information. *Quarterly Journal of Experimental Psychology*, 1959, *11*, 92-107.

Wason, P.C. Response to affirmative and negative binary statements. *British Journal of Psychology*, 1961, *52*, 133-142.

Williams, A.R., Jr., Siegel, A.I., & Burkett, J.R. *Readability of textual materials--A survey of the literature.* (AFHRL-TR-74-29) AD-785140. Lowry AFB., Colo.: Technical Training Division, Air Force Human Resources Laboratory, July 1974. Also: (MS. No. 876) JSAS *Catalog of Selected Documents in Psychology*, 1975, *5*, 202.

Williams, A.R., Jr., Siegel, A.I., Burkett, J.R., & Groff, S.D. Final report of Applied Psychological Services Project Name *MERIT* Lowry AFB., Colo.: Technical Training Division Air Force Human Resources Laboratory, in press.

Winer, B.J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1962.

Yngve, V.R. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, 1960, *104*, 444-466.

APPENDIX A

Summary Specifications for Each Measure

COGNITION OF SEMANTIC UNITS

1. Measure number — 1

2. Category — Structure-of-Intellect

3. Abbreviation — CMU

4. Explanation — Type/Token ratio. A function of the number of different words in a block of text, NDWB, and the total number of words in the block, TNWB

5. Computational complexity — Readily automated

6. Dictionary requirements — Entries will include abbreviations as well as words. Each will be so identified.

7. Symbolic definition

$$1 - \frac{NDWB}{TNWB} = CMUB$$

8. Scaling — $0 < CMUB \leq 1$
1= most comprehensible

9. Rules utilized

(1)A word is defined to consist of any number of consecutive alphanumerical characters preceded by a space and followed by a comma, space, exclamation point, colon, or question mark.

(2)Two words are the same only if they are spelled exactly the same (i.e., prefixes, tenses, plurals, etc., will be taken into account and the word "walk" and "walked" will be counted as different words).

(3)Abbreviations of multiple words (e.g., "USAF," "USSR," and "APA" will each be counted as one word. (A count of the number of such abbreviations, NAMB, determined in calculating CMU will be retained for later use in calqulating ESI.)

(4)Hyphenated words will be counted as multiple words (i.e., a hyphen will be considered like a space) except in the following prefixes and suffixes and post fixes:

pre-war
post-war

(e.g., "never-to-be-forgotten" is 4 words) This same rule applies to words containing a slash such as "upper/lower."

(5) Each numerical value will be counted as one word (e.g., 4.56, 1×10^{-3} or 4π).

(6) Capitalization of the first letter of a word will be ignored. Thus, a word which is capitalized because it starts a sentence or appears in a title is counted as the same word as if it had been composed of all lower case letters which appear elsewhere. (This is not true when the word ends in a period - see 13 below.)

(7) A word composed entirely of capital letters is counted as different from the same word in lower case letters.

(8) Italics are ignored in word counting; thus, an italicized word is counted as the same word as one not italicized.

(9) One-character symbols which occur as one-symbol words will be counted as one word.

(10) Each word in a spelled out number will be counted as one word (e.g., "eight hundred" will be counted as two words, whereas "800," from above, is counted as one word.

(11) Tables, figures, maps, diagrams, illustrations and the like are not considered in this measure, but titles of these are included.

(12) Words within titles and headers will not be included in counts of NDWB and TNWB.

(13) All one-word abbreviations composed of any combination of upper and lower case alphanumeric letters will be counted as one word. An abbreviation here is defined to be any word with or without interspersed periods followed by a period. (The number of such single word abbreviations, NASB, is retained for later use in determining ESI.

(14) A sentence will be determined by scanning text and identifying groups of words such that:

a. the first starts with an initial capital letter
b. the last word ends with a period, question mark or exclamation point, and is followed by one or more spaces and then a word with an initial capital letter.

COGNITION OF SEMANTIC RELATIONS

1. Measure number — 2

2. Category — Structure-of-Intellect

3. Abbreviation — CMR

4. Explanation — Number of shared nouns in a text block NSNB per sentence pair plus the number of references in a text block NORB, divided by the number of words in the block TNWB.

   This measures the number of links or relationships in a text block.

5. Computational complexity — Moderate

6. Dictionary requirements — Identify reference words

7. Symbolic definition —

$$CMRB = \frac{\left(\frac{NSNB}{TNSB - 1}\right) + NORB}{TNWB}$$

8. Scaling — Highest value = most comprehensible
   $0 \leq CMRB \leq 1$

9. Rules utilized — (1) The number of references NORB is determined by counting the total number of pronouns. This will require sentence parsing to determine which words are used as pronouns.

   (2) The number of shared nouns NSNB is calculated by considering adjacent sentence pairs and tallying the number of times per pair a common noun occurs in both sentences (plural nouns are considered as if they were singular in this process):

COGNITION OF SEMANTIC RELATIONS (cont.)

Sentence	Pairs	No. of nouns appearing in both sentences of pair
1		n_1
2		n_2
3		n_3
4		n_4
.		.
.		.
.		.
i		n_i
.		.
.		.
.		.
N		

$$\text{Calculate: } \frac{\sum_{i=1}^{N-1} n_i}{N-1} = \frac{NSNB}{TNSB - 1}$$

This will require sentence parsing in which each word is identified as a noun or not. Words which are ambiguous (i.e., could be noun or another part of speech and parser cannot distinguish) will be considered as nouns. Multiple occurrences of a single noun are counted as one occurrence, e.g., a given noun appearing twice in one sentence and three times in the adjacent sentence is tallied as a single occurrence in determining n_i. No test will be made on the meaning of the nouns, and if the words match it will be assumed they have the same meaning.

(3) TNWB is the same as was determined in CMU.

MEMORY OF SEMANTIC UNITS

1.	Measure number	3
2.	Category	Structure-of-Intellect
3.	Abbreviation	MMU
4.	Explanation	This measure is a function of the number of different nouns per block of text, NDNB and the total number of words in the block TNWB
5.	Computational complexity	Requires sentence parsing for determination of nouns.
6.	Dictionary requirements	Identify parts of speech in all words including all one-syllable symbols.
7.	Symbolic definition	$MMUB = 1 - \frac{NDNB}{TNWB}$
8.	Scaling	$0 < MMUB < 1$ 1= most comprehensible
9.	Rules utilized	(1) Words will be identified as nouns by a dictionary search and additional parsing as required. (2) Two nouns are the same only if they are spelled exactly the same (i.e., "grass" and "grasses" will be counted as different words. (3) Abbreviations will be counted as no more than a single noun (e.g., "USA," S.P.C.A., and "Soc." are single nouns). Some abbreviations, not referring to or containing a noun will not be counted as a noun (e.g., i.e.,). This will be determined via a dictionary search. (4) The following rules from CMU apply: 4- words in with hyphens, slashes 6 - initial caps ignored 7 - all caps 8 - italics ignored 12 - tables, figures ignored 13 - titles, headers excluded

(5) the number of nouns to be counted in symbols will be contained in the dictionary, e.g.,

¢ 1

$\leq$ 0

(6) A single noun appearing twice within a block with different meanings will count as 1 in NDNB.

(7) Any numerical value which begins with the symbol $ or £ or ends with ¢ or % will not be counted as nouns.

EVALUATION OF SYMBOLIC IMPLICATIONS

1.	Measure	4
2.	Category	Structure-of-Intellect
3.	Abbreviation	ESI
4.	Explanation	This measure is a function of the number of abbreviated or symbolic words in a block of text, NSWB, and the total number of words in the block, TNWB.
5.	Computational complexity	Readily automated
6.	Dictionary requirements	No unique requirements
6.		
7.	Symbolic definition	$1 - \frac{NSWB}{TNWB} = ESIB$
8.	Scaling	$0 \leq CMUB \leq 1$ 1= most comprehensible
9.	Rules utilized	(1) The number of abbreviated or symbolic words NSW is equal to the sum of a. no. of multiple word abbreviations, NAM (see item 3 of CMU) b. no. of single word abbreviations, NAS (see item 14 of CMU) c. no. of one character symbols (see item 9 of CMU) (2) TNWB is as calculated in CMU. (3) Numbers with or without decimal points do not count as abbreviations. (4) A word starting with \$ and ending with ¢ counts as a single word.

CONVERGENT PRODUCTION OF SEMANTIC IMPLICATIONS

1.	Measure number	7
2.	Category	Structure-of-Intellect
3.	Abbreviation	NMI
4.	Explanation	Measure of the average number of parts of speech per word in a text block. Measures frequency of need for reader to make inferences.
5.	Computational complexity	Readily automated
6.	Dictionary requirements	Parts of speech for each word
7.	Symbolic definition	$NMIB = \frac{TNWB}{TPSB}$ Where TPSB is the total number (not total different number) of parts of speech in all words in a text block.
8.	Scaling	$0 \leq NMIB \leq 1$ highest is most comprehensible
9.	Rules utilized	(1) Count past participle as 2 parts of speech, present participles as 3.

DIVERGENT PRODUCTION OF SEMANTIC UNITS

1.	Measure number	8
2.	Category	Structure-of-Intellect
3.	Abbreviation	DMU
4.	Explanation	Measure of the number of elucidations or explanations per sentence in a block of text= NESS.
	(More desirable, but not considered would be	% of need for explanations filled)
5.	Computational complexity	Modest
6.	Dictionary requirements	Store file of key words which identify that an explanation is forthcoming.
7.	Symbolic definition	$DMUB = \frac{\sum_S NESS}{TNSB} = \frac{TNEB}{TNSB}$
8.	Scaling	$0 \leq DMUB \leq 1$ highest= most comprehensible
9.	Rules utilized	(1) Count one explanation for each occurrence of the following word or word combinations: that is i.e. thus consequently in other words therefore to illustrate for example No more than one occurrence of an explanation is counted per sentence.

DIVERGENT PRODUCTION OF SEMANTIC UNITS (cont.)

(2) Under conditions to be specified, occurrence of the following words (if used in connection with an explanation) will be counted as an explanation:

elucidate
explain
illustrate, illustration
expound
instance
case
example

YNGVE DEPTH

1. Measure number — 9

2. Category — Psycholingusitic

3. Abbreviation — YD

4. Explanation — In order to determine the depth or complexity of a sentence as defined by Yngve, lines coming out of each node of a parse diagram are numbered 0, 1,... from right to left. Each word W, in the sentence (at the bottom of the diagram) is assigned a value equal to the sum of the numbers along the path from initial symbol (S) to that word. The Yngve depth of a given word in the sentence, YD(W) is the average of these numbers.

5. Computational complexity — Difficult since it is based upon need for complete sentence parsing.

6. Dictionary requirements — Parts of speech as required for parsing.

7. Symbolic definition

$$YD \begin{cases} YD(W) = \text{sum of all digits on parse path to the given word} \\ YDS = \dfrac{TNWS}{\sum_{W} YD(W)} \\ YDB = \dfrac{\sum_{S} YDS}{TNSB} \end{cases}$$

8. Scaling — $0 \leq YDB \leq 1$
Higher YDs correspond to most comprehensible sentences.

9. Rules utilized — (1) Parse sentence.
(2) Number the lines coming out of each node 0, 1, ... from right to left. Assign to each word the sum of the numbers along the path from the initial symbol to the word. YD is the average of these numbers.

(3) Do not include title, headers.

(4) If the sentence is ambiguous, i.e., has more than one potential parse, then the parse having the largest value of YDS is selected.

(5) A limit to the number of potential parses to be sought by the computer will be specified as input.

MORPHEME DEPTH

1. Measure number — 10

2. Category — Psycholinguistic

3. Abbreviation — MD

4. Explanation — Count of the number of words per text block, divided by the number of morphemes in the block, TNMB.

5. Computational complexity — Modest

6. Dictionary requirements — Store number of morphemes for each dictionary entry

7. Symbolic definition — $MDB = \frac{TNWB}{TNMB}$

8. Scaling — $0 \leq MDB \leq 1$
   Highest is most comprehensible

9. Rules utilized — (1) Find morpheme counts per word via dictionary lookup.

   (2) Each numerical value is counted as one morpheme.

   (3) All abbreviations, whether one word (Mr) or multiple words (USAF) are tallied as a single morpheme.

   (4) Capitalization is ignored in morphe counting.

   (5) A one-character symbol which occurs as one symbol word is counted as one morpheme.

   (6) Clichés will be counted as one morpheme.

TRANSFORMATIONAL COMPLEXITY

1. Measure number — 11

2. Category — Psycholinguistic

3. Abbreviation — TC

4. Explanation — Determination as to whether each clause is active, passive, passive-negative, or active-negative.

5. Computational complexity — Less than that or parsing, as method depends on partial output of parser.

6. Dictionary requirements — Parts of the verb "to be" must be labeled as such; verbs must be labeled as transitive or intransitive; past participles of verbs must be labeled.

7. Symbolic definition

$$TCS = \begin{cases} 1.00 \text{ if sentence is active} \\ 0.95 \text{ if sentence is passive} \\ 0.75 \text{ if sentence is active-negative} \\ 0.20 \text{ if sentence is passive-negative} \end{cases}$$

$$TCB = \frac{\sum_{S} TCS}{TNCB}$$

$$TNCB = \frac{TNCS}{TNSB}$$

8. Scaling — $0.2 \leq TCB \leq 1.0$
Higher numbers correspond to sentences which are easier to comprehend.

9. Rules utilized — (1) Isolate the complex verb constituent of the main clause.

   (2) If the last two words of this structure are:
   - (a) the past, present, future or infinitive of the verb "to be": followed by
   - (b) the past participle of a transitive verb then the sentence is in the passive voice.

(3) Otherwise, it is in the active voice.

(4) If there is one occurrence of any form of the word "not" within this complex verb constituent, then the sentence is considered negative. These include:

never
not
neither
none
words ending with n't

(5) Thus, if the sentence is both passive and negative, it is called passive-negative.

(6) If a sentence is ambiguous, and the different meanings of it lead to different values for TC, then omit the measure for that sentence.

(7) Each clause must have a subject and predictate.

CENTER EMBEDDING

1. Measure number — 12

2. Category — Psycholinguistic

3. Abbreviation — CE

4. Explanation — A measure of the number of chained modifying clauses on the right of the subject noun phrase of a sentence.

5. Computational complexity — Equal to that of parsing

6. Dictionary requirements — None beyond those of parsing

7. Symbolic definition — The number of sentences in a block, TNSB, divided by the number of phrases to the right of the subject noun in a sentence, NNPS, summed over all sentences in a block:

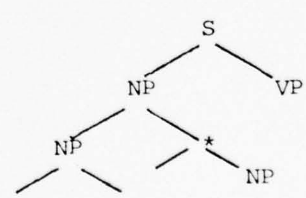
$$CEB = 1 - \frac{\sum_{S} NNPS}{TNSB}$$

(The NNP of NP VP is 0.)

8. Scaling — $0 \leq CEB \leq 1$
The higher the SE, the easier the sentence.

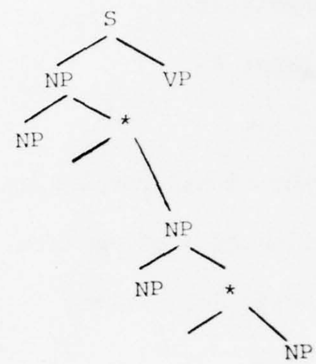
9. Rules utilized — (1) Parse the sentence.

(2) If the parse is of the following form, NNPS= 1:



*Any sentence element (category)

(3) If the parse is of the following form, NPPS= 2:



*Any sentence element (category)

RIGHT BRANCHING

1. Measure number — 13

2. Category — Psycholinguistic

3. Abbreviation — RB

4. Explanation — A measure of the number of chained modifying clauses (elaborations) on the right of the object noun phrase of a sentence.

5. Computational complexity — Equal to that of parsing.

6. Dictionary requirements — None beyond those of parsing

7. Symbolic definition — The number of sentences per block, TNSB, divided by the number of chained modifying clauses on the right of the object noun phrase, NCRS, summed over all sentences in a block.

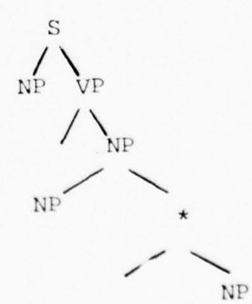
$$RBB = 1 - \frac{\sum_{S} NCRS}{\sum_{S} NCRS + TNSB} = \frac{TNSB}{\sum_{S} NCRS + TNSB}$$

(The NCSR of NP VP is 0.)

8. Scaling — $0 \leq ROB \leq 1$
The higher the RB, the more comprehensible the sentence.

9. Rules utilized — (1) Parse the sentence.

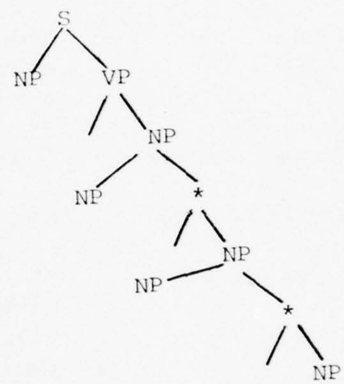
(2) If the parse is of the following form, NCRS= 1:



*Any sentence element (category)

RIGHT BRANCHING (cont.)

(3) If the parse is of the following form, NCRS= 2:



LEFT BRANCHING

1.	Measure	14
2.	Category	Psycholinguistic
3.	Abbreviation	LB
4.	Explanation	A measure of the number of chained modifying clauses on the left of the subject noun of a sentence.
5.	Computational complexity	Equal to that of parsing.
6.	Dictionary requirements	None beyond that of parsing.
7.	Symbolic definition	The number of sentences per block, TNSB, divided by the number of chained modifying clauses on the left of the subject noun of a sentence, NCLS, summed over all sentences in a block:

$$LBB = 1 - \frac{\sum_S NCLS}{TNSB}$$

(The NCLS of NP VP is 0.)

8. Scaling

$0 \leq LBB < 1$
The higher the LB, the more comprehensible is the sentence.

9. Rules utilized

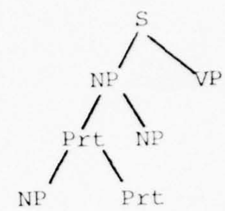
(1) Parse the sentence.

(2) Count the number of modifying clauses.

Left branching involves the presence of clauses modifying the subject noun phrase of a sentence, as in:

The stogie-chewing dictator laughed.

The presence of such a clause requires that the parse be of the following form:



The newly introduced noun phrase can be itself rewritten in terms of a new noun phrase and a participle phrase, to yield sentences such as:

> *The smoke-emitting-stogie-chewing dictator laughed.*

Forms of the type:

> *The belly-laughing stogie-smoking dictator laughed.*

are not of the left-branching variety because *belly-laughing* does not modify stogie.

<u>COUNT</u>: A sentence without the left-branching feature being described has NCLS= 0. Every time there is a rewriting of the leading noun phrase as a participle phrase and noun phrase, add 1 to NCLS.

DELETED COMPLEMENT

1. Measure number — 15

2. Category — Psycholinguistic

3. Abbreviation — DC

4. Explanation — Measure of frequency with which the complementizer "that" of a noun phrase complement has been deleted e.g., as in: "the fact he came was not obvious."

5. Computational complexity — Equal to that of parsing.

6. Dictionary requirements — Complementizers must be labeled as such.

7. Symbolic definition — DCS= 0 if complementizer is present
DCS= 1 if relative pronoun is deleted.

$$DCB = 1 - \frac{\sum_{S} DCS}{TNSB}$$

8. Scaling — $0 \leq DCB \leq 1$
Higher DC corresponds to more comprehensible sentences.

9. Rules utilized — (1) Parse the sentence.

(2) If at any point in the parsing, a noun phrase complement with deleted complementizer is present, then DCS= 0.

(3) Otherwise, DCS= 1.

APPENDIX B

<u>Detailed Specifications for Each Program Module</u>

The regression equations presented in this section should be regarded as approximate. These equations for predicting comprehensibility for high reading grade level readers, low reading grade level readers, and for both types combined are preliminary and subject to modification as additional work with these equations unfolds.

NAME: INITIAL

NUMBER: 1

PURPOSE: The INITIAL module performs all initialization needed to begin a run.

TECHNIQUE: The INITIAL module determines whether the request input is coming from cards or remote terminal and initializes accordingly. All other startup functions requred are performed at this time. See Appendix C.

INPUT: None

FILES ACCESSED: Sets up INPUTFILE.

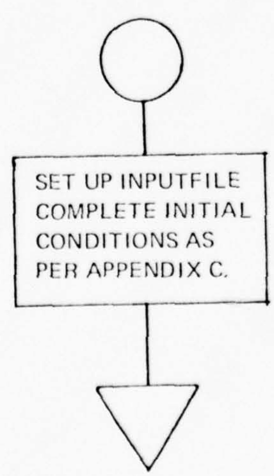
GLOBAL DATA: None

OUTPUT: The output of the INITIAL module is the ability to begin measure processing.

MODULES CALLED: None

CALLING MODULES: None

COMMENTS: The exact functions necessary in INITIAL depend on the particular implementation system.



NAME: READ

NUMBER: 2

PURPOSE: The READ module processes one input request via SCANINPUT. The input request is syntaxed. If the request is valid, the READ module sets up any default (unspecified) internal values and all specified values. If syntax errors are detected, the request is ignored and all errors reported.

TECHNIQUE: READ calls on SCANINPUT to obtain the first word of the request. If the first word is not "CHECK" or "MEASURE," an error is reported and the request is ignored. All of the default values are set up for either a CHECK or MEASURE request. The major portion of READ is a "standard" table-driven syntax analysis routine. Each request item is identified by its keyword. READ checks that no individual specification is provided more than once.

If one or more syntax errors are discovered, READ will report each error to the user via ERROR. If possible, the remainder of that request will be syntaxed in order to provide the maximum information to the user. If direct syntaxing can not be continued, READ will scan the input request (using SCANINPUT) looking for the next key word. Syntax checking will resume when a key word is found. If a period is found (ending the current request), READ will stop.

When the period ending the request is scanned, READ will exit if there were no syntax errors or will start over (to process the next request) if there were syntax errors.

INPUT: One input run request ("CHECK...." or "MEASURE....") contained on one or more input records (cards or remote transmissions).

FILES ACCESSED: INPUTFILE (via SCANINPUT): card or remote file containing input request.
TEXTFILE: to verify presence and validity.
DICTFILE: to verify presence and validity.
OUTPUTFILE: to report errors (via ERROR) and/or report type of request being performed. (SCANINPUT will provide here a list of the input records read.)
EXAMPLEFILE: to verify presence and validity.
CLICHESFILE: to verify presence and validity.

GLOBAL
DATA: Sets up:

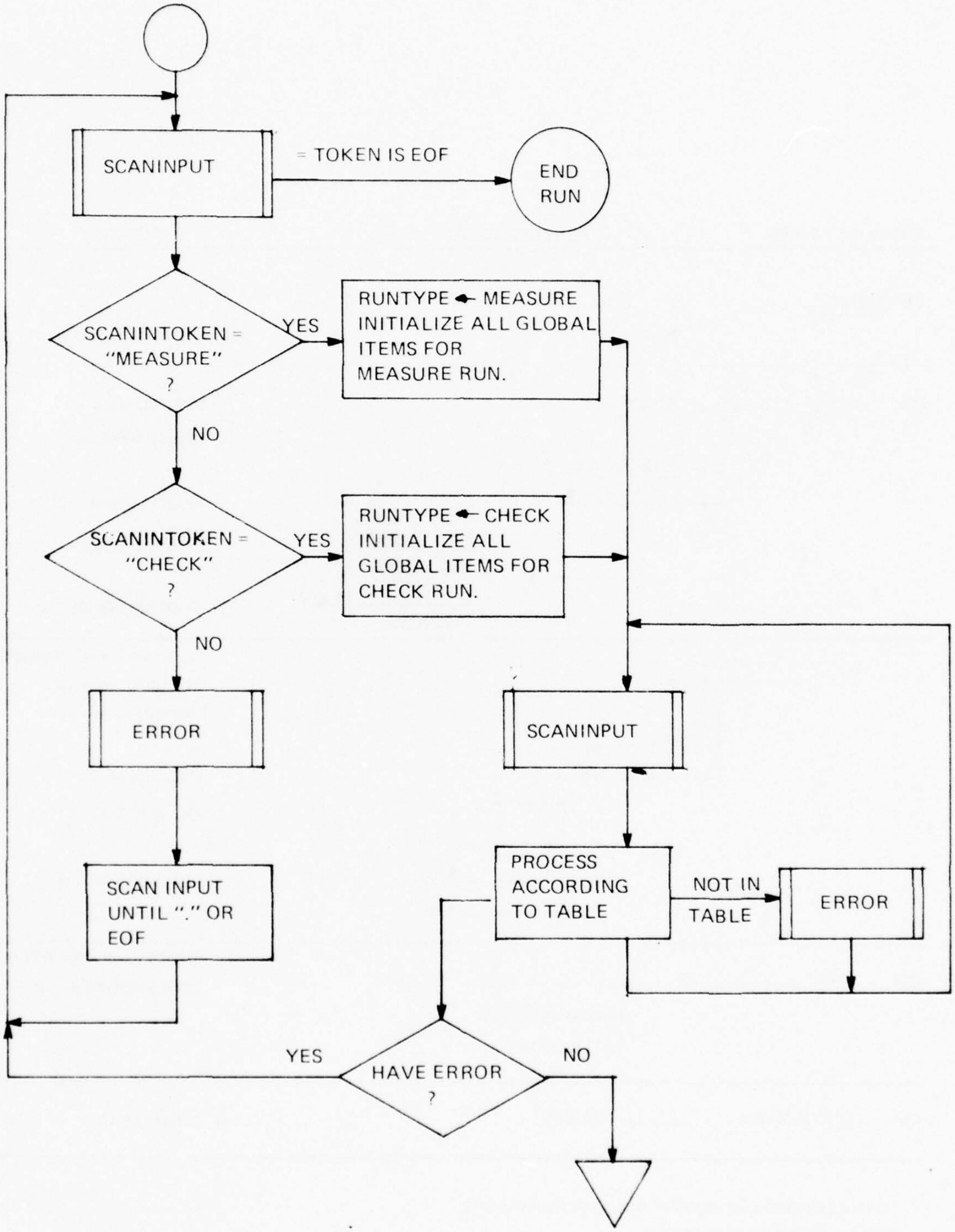
RUNTYPE
DICTFILE title & media, verify presence
TEXTFILE title & media, verify presence
EXAMPLEFILE title & media, verify presence
CLICHEFILE title & media, verify presence
STARTLINE, STARTWORD
ENDLINE, ENDWORD
BLOCKTYPE, BLOCKSIZE
LISTLOC, LISTLINE, LISTTEXT
OUTPUTFILE media
MEASURE [*]
MODE
SAMPLESIZE (only if CHECK).
ABORTPERCENT
NORMCDC, NORMMAN, NORMTO, NORMSG, NORMOVERALL, READERHIGH, READERLOW
COMMENT
MAXPARSE
BLOCKCOUNT
ENDTYPE

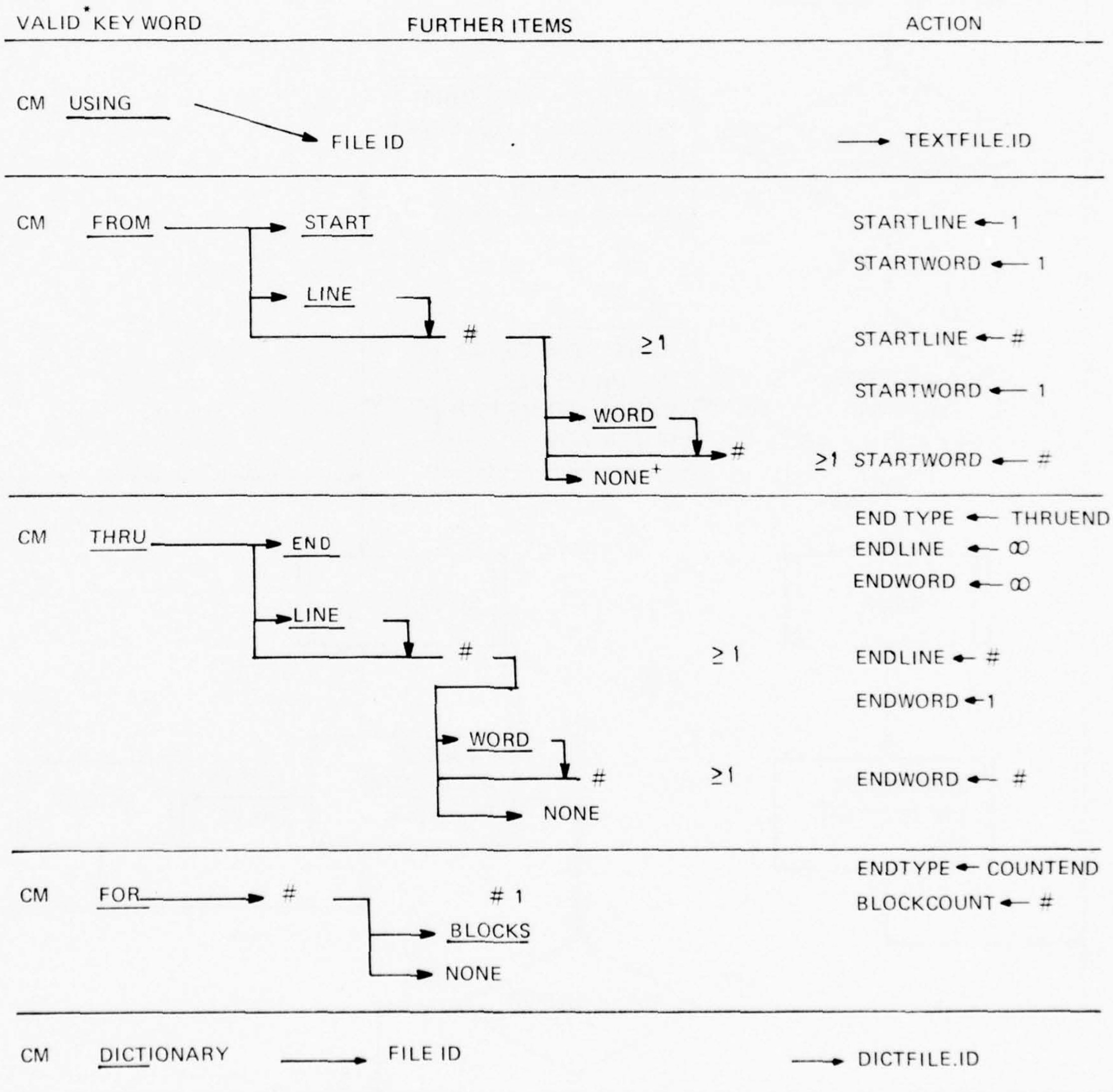
OUTPUT: The output of the READ module is the initial values of the global data as described above.

MODULES
CALLED: SCANINPUT for next input token.
ERROR to report error.

CALLING
MODULES: None

COMMENTS: The size of the READ module is dependent on the amount of error checking to be performed. It will likely be the second largest module (after PARSE) but will contain no new or difficult concepts. It can be built quickly; initially it can serve as an aid to other module debugging but should have complete error checking and error messages when released to the nontechnical user.


SCANINPUT
= TOKEN IS EOF
END RUN
SCANINTOKEN = "MEASURE" ?
YES
RUNTYPE ← MEASURE INITIALIZE ALL GLOBAL ITEMS FOR MEASURE RUN.
NO
SCANINTOKEN = "CHECK" ?
YES
RUNTYPE ← CHECK INITIALIZE ALL GLOBAL ITEMS FOR CHECK RUN.
NO
ERROR
SCANINPUT
SCAN INPUT UNTIL "." OR EOF
PROCESS ACCORDING TO TABLE
NOT IN TABLE
ERROR
YES
HAVE ERROR ?
NO



* "C" VALID FOR CHECK; NOT VALID FOR MEASURE

\+ NONE – NONE OF THE ABOVE

CM BLOCK → MARKS — BLOCKTYPE ← CHAR.

SIZE → # — ≥ 1 — BLOCKSIZE ← #

→ SENTENCES — BLOCKTYPE ← SENTENCES

→ WORDS — BLOCKTYPE ← WORDS

→ NONE — BLOCKTYPE ← WORDS

CM LIST → TO

→ PRINTER — LISTLOC ← PRINTER

→ TERMINAL — LISTLOC ← REMOTE

M TEXT — LISTTEXT ← TRUE

→ WITH → LINE → NUMBERS — LISTLINE ← TRUE

→ NONE

M RGL — SET CORRES. MEASURE [•]

M COMPREHENSIBILITY → MEASURE — SUMMARY

→ SUMMARY

→ DETAIL

→ NONE

SET CORRESPONDING MEASURE [•]

M <TEST NAME> — SET CORRES. MEASURE [•]

CM	BATCH		MODE ← 3ATCH
CM	INTERACTIVE		MODE ← INTERACTIVE
C	EVERY	→ # → WORDS / NONE	SAMPLESIZE ← #
CM	ABORT	→ AT / ON → # → % / NONE	ABORT PERCENT ← #
M	EXAMPLE	→ FILE ID	EXAMPLE FILE.ID ← FILEID
M	CLICHE	→ FILE ID	→ CLICHESFILE.ID ← FILEID
M	PARSE	→ MAX → #	MAXPARSE ← #
CM	COMMENT	<QUOTED STRING>	← COMMENT <QUOTED STRING>
M	NORM	→ MANUALS / CDC / TO / SG / OVERALL / NONE	NORMOVERALL ← FALSE NORMMAN ← TRUE NORMCDC ← TRUE NORMTO ← TRUE NORMSG ← TRUE NORMOVERALL ← TRUE
CM			
M	READER	→ HIGH / HIGHLOW / LOW / NONE	READERHIGH ← READERLOW ← READERHILOW ← FALSE READERHIGH ← TRUE READERHILOW ← TRUE READERLOW ← TRUE

NAME: SCAN

NUMBER: 3

PURPOSE: The SCAN module reads the text file and finds the next block to process. SCAN also sets up all global items necessary to process a new block.

TECHNIQUE: When SCAN is called the first time (BLOCKNUMBER= 0), SCAN reads the text file looking for word STARTWORD on line STARTLINE. For all other calls (BLOCKNUMBER > 0), the text file is already positioned immediately before the first word of the next block. In either case, all block-level global values are set up for a new block and BLOCKNUMBER is increased by one.

INPUT: None

FILES ACCESSED: TEXTFILE

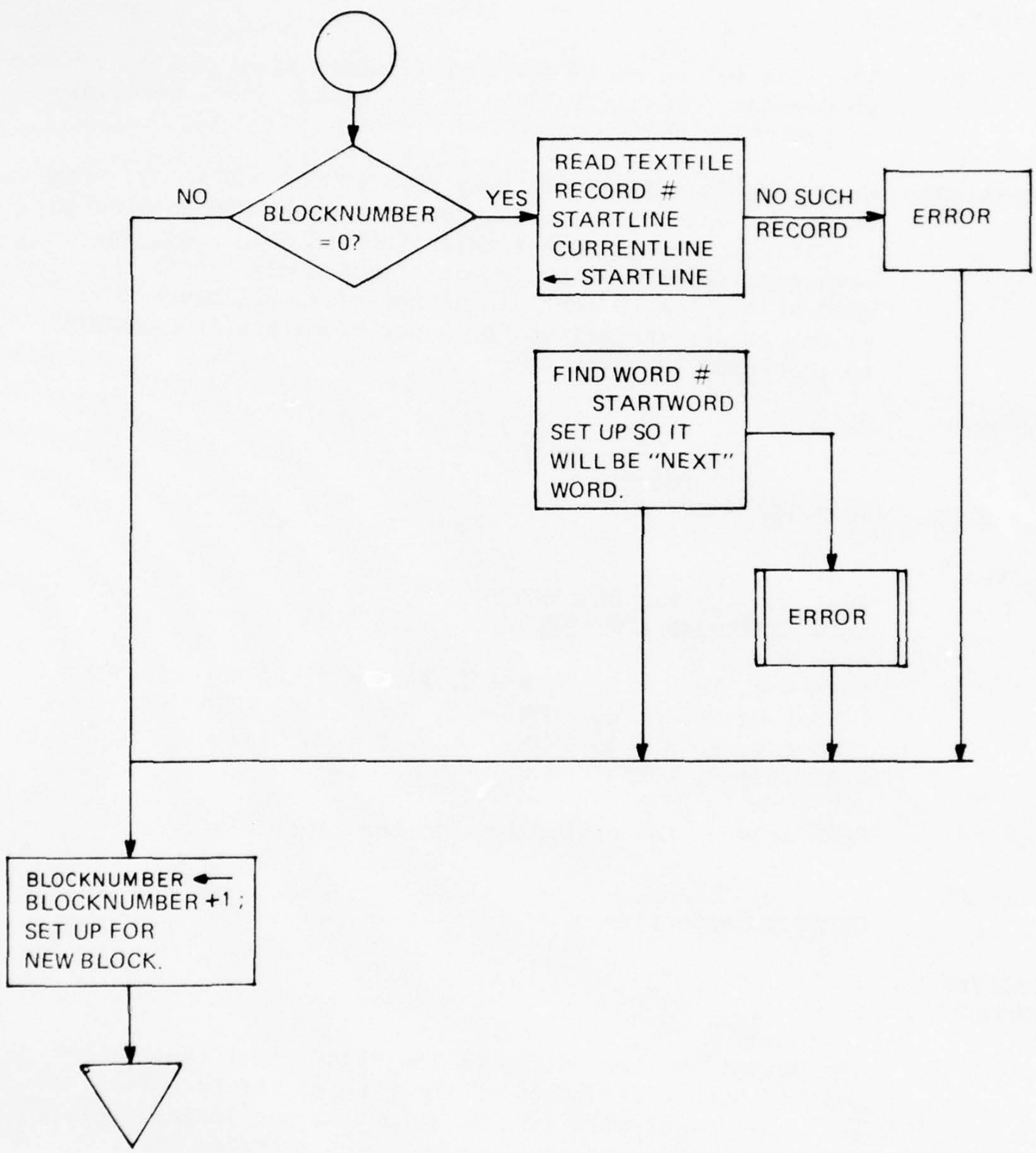
GLOBAL DATA: uses and updates BLOCKNUMBER
uses STARTWORD & STARTLINE
sets up:
NDWB ← 0, TNWB ← 0, NAMB ← 0, NASB ← 0, NSNB ← 0, NDNB ← 0, NSWB ← 0, TNSE ← 0, TNMB ← 0, NPPB ← 0, NWNDB ← 0, TPSB ← 0, TNCB ← 0, NORB ← 0, TNEB ← 0, TCLB ← 0, OSWB ← 0, TSCB ← 0, NDNSAVE ← 0, NDWSAVE ← 0.

OUTPUT: SCAN sets up the system for the next text block.

MODULES CALLED: ERROR to report error.

CALLING MODULES; None

COMMENT: The actual implementation of the "find first text block" part of SCAN will depend on the form of the text file. It is assumed that each line is in a separate logical record and that logical records are numbered starting at 1.


NO
BLOCKNUMBER = 0?
YES
READ TEXTFILE RECORD # STARTLINE CURRENTLINE ← STARTLINE
NO SUCH RECORD
ERROR
FIND WORD # STARTWORD SET UP SO IT WILL BE "NEXT" WORD.
ERROR
BLOCKNUMBER ← BLOCKNUMBER +1 ; SET UP FOR NEW BLOCK.

NAME: SCANINPUT TYPE= REAL

NUMBER: 4

PURPOSE: SCANINPUT is the only module which directly accesses user input. Each time SCANINPUT is called, it returns the next token in the user input. Whenever necessary it will read the next input record (card or remote).

TECHNIQUE: The exact method used by SCANINPUT is dependent on the particular language chosen. A token in the user input is defined as one of:

1. a string of letters terminated by a blank or end of input record or a non-letter. (A letter is A,B,C,D,E,...,Z.) This type of token is called a word.
2. a string of digits (0,1,2,...,9) terminated by a blank or end of input record or a non-digit. This type of token is called a number.
3. a single character which is not a letter, digit, or blank, called a special symbol.
4. a < file id > (in the form specified by the particular implementation), called a file id.

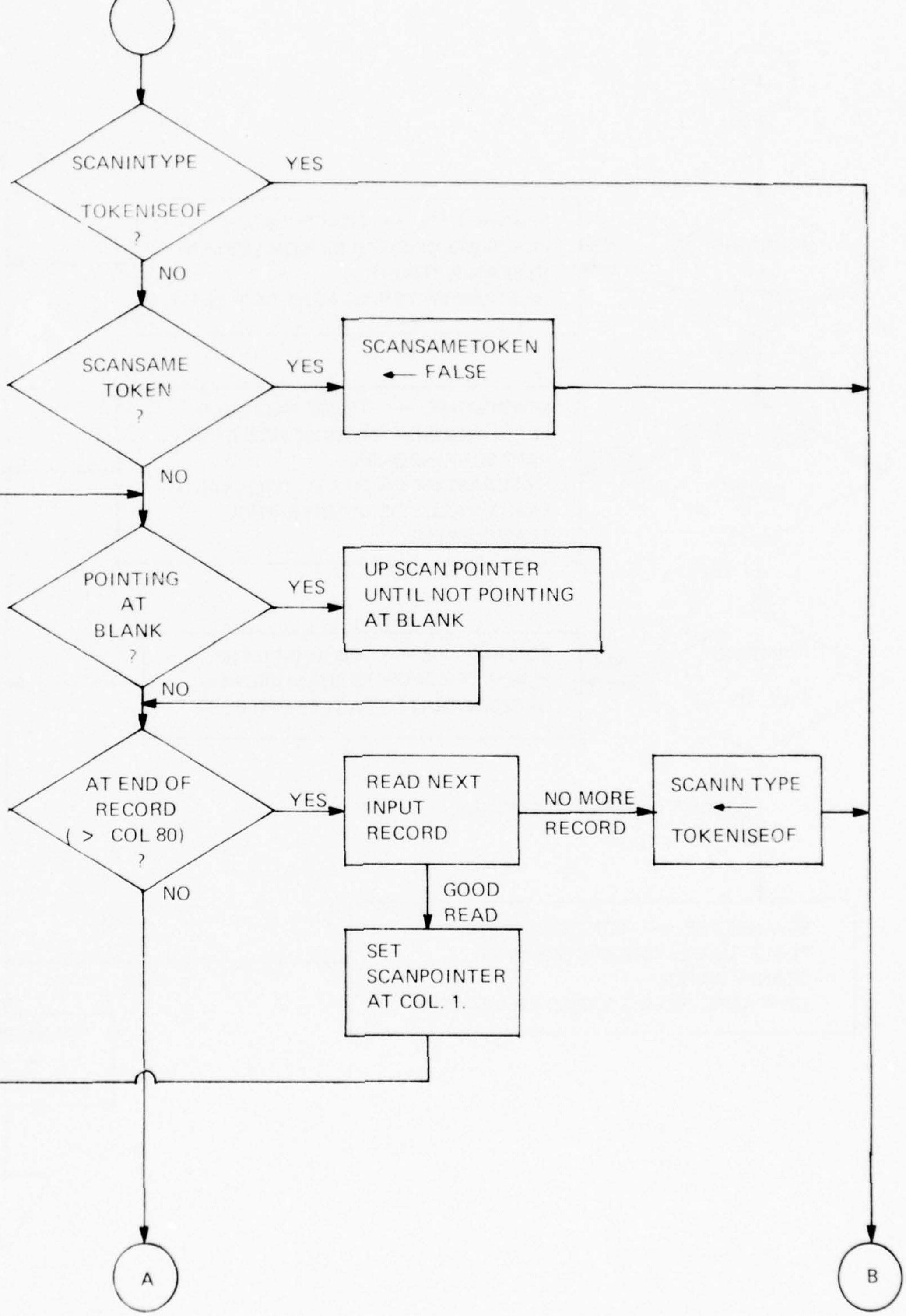
INPUT: SCANRECORD - current input record.
SCANPOINTER - current place in SCANRECORD.
SCANSAMETOKEN - if TRUE, SCANINPUT will not look at the next token but will re-look at the same token.

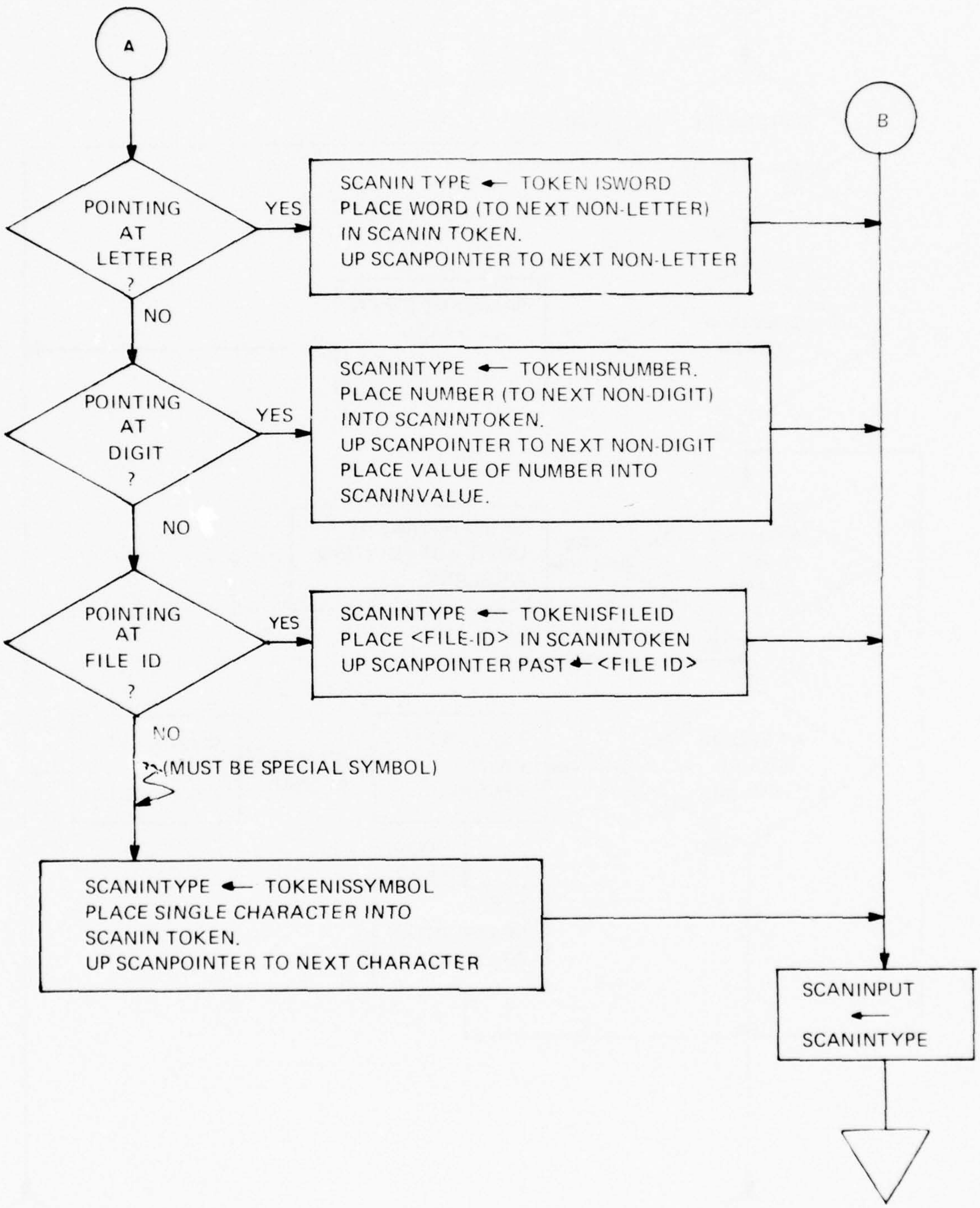
FILES ACCESSED: INPUTFILE.

GLOBAL DATA: SCANRECORD (changed if a new input record needed).
SCANPOINTER (updated unless SCANSAMETOKEN is TRUE).
SCANINTYPE set up.
SCANINVALUE set up if token is a number.
SCANSAMETOKEN ← FALSE.
SCANINTOKEN set up.

OUTPUT: The output of SCANINPUT is a token from the user input. The token itself is placed in SCANINTOKEN. Its type is placed in SCANINTYPE and is also returned as the value of the module. If the token is a number, SCANINVALUE contains the integral value represented by the token. For example, if SCANINTOKEN is "0132" then SCANINVALUE= 132.

MODULES CALLED:	None
CALLING MODULES:	READ SEARCH
COMMENTS:	In general, SCANINPUT is an easy module to write. Difficulty can be introduced by the choice of language or system. SCANINPUT can also print each input line as read.


SCANINTYPE
TOKENISEOF
?
YES
NO
SCANSAME
TOKEN
?
YES
SCANSAMETOKEN
← FALSE
NO
POINTING
AT
BLANK
?
YES
UP SCAN POINTER
UNTIL NOT POINTING
AT BLANK
NO
AT END OF
RECORD
(> COL 80)
?
YES
READ NEXT
INPUT
RECORD
NO MORE
RECORD
SCANIN TYPE
←
TOKENISEOF
GOOD
READ
SET
SCANPOINTER
AT COL. 1.
NO
A
B


A
B
POINTING AT LETTER ?
YES
SCANIN TYPE ← TOKEN ISWORD
PLACE WORD (TO NEXT NON-LETTER)
IN SCANIN TOKEN.
UP SCANPOINTER TO NEXT NON-LETTER
NO
POINTING AT DIGIT ?
YES
SCANINTYPE ← TOKENISNUMBER.
PLACE NUMBER (TO NEXT NON-DIGIT)
INTO SCANINTOKEN.
UP SCANPOINTER TO NEXT NON-DIGIT
PLACE VALUE OF NUMBER INTO
SCANINVALUE.
NO
POINTING AT FILE ID ?
YES
SCANINTYPE ← TOKENISFILEID
PLACE <FILE-ID> IN SCANINTOKEN
UP SCANPOINTER PAST ← <FILE ID>
NO
(MUST BE SPECIAL SYMBOL)
SCANINTYPE ← TOKENISSYMBOL
PLACE SINGLE CHARACTER INTO
SCANIN TOKEN.
UP SCANPOINTER TO NEXT CHARACTER
SCANINPUT
←
SCANINTYPE

NAME: ERROR ERROR (ERRNUM)

NUMBER: 5

PURPOSE: ERROR reports to the user a syntax in the user's input.

TECHNIQUE: ERROR simply reports the specified error message (as indicated by ERRNUM) including the last scanned input token. ERROR also indicates the presence of a syntax error by adding 1 to ERRORCOUNT and setting HAVEERROR.

INPUT: ERRNUM (real internal error number).

FILES ACCESSED: OUTPUTFILE

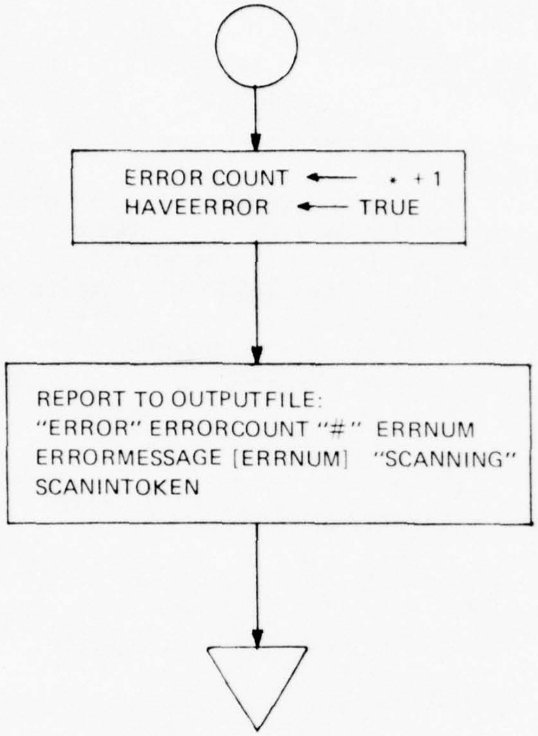
GLOBAL DATA: ERRORCOUNT ← ERRORCOUNT + 1
HAVEERROR ← TRUE

OUTPUR: The only output from ERROR is the error message to the user.

MODULES CALLED: None

CALLING MODULES: READ SCAN
SEARCH

COMMENTS: ERROR is a very simple and very short module containing a list of error messages. The list (only) is increased with the complexity of the syntax checking routines of READ & SEARCH.


ERROR COUNT ← • + 1
HAVEERROR ← TRUE
REPORT TO OUTPUTFILE:
"ERROR" ERRORCOUNT "#" ERRNUM
ERRORMESSAGE [ERRNUM] "SCANNING"
SCANINTOKEN

NAME: RESET

NUMBER: 6

PURPOSE: The RESET module sets up all global items necessary to process a new sentence.

TECHNIQUE: All sentence counters are reset and other sentence-level global items are cleared.

INPUT: None

FILES ACCESSED: None

GLOBAL DATA: Sets up:

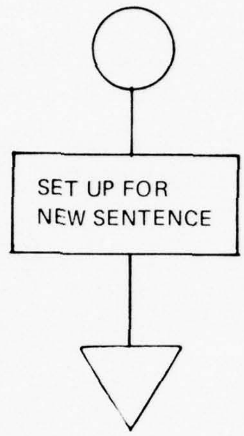
TNSB ← TNSB + 1
NPPS ← 0
NESS ← 0
TNWS ← 0
YDS ← 0
TCS ← 0
NNPS ← 0
NCRS ← 0
NCLS ← 0
DCS ← 0

OUTPUT: RESET sets up the system for the next sentence.

MODULES CALLED: None

CALLING MODULES: None

COMMENTS: Few problems, if any, are associated with this module.


SET UP FOR
NEW SENTENCE

NAME: SEARCH

NUMBER: 7

PURPOSE: The SEARCH module looks up a single text word in the dictionary and adds the dictionary information for that word to the Sentence Data Array.

TECHNIQUE: SEARCH adds one to TNWS (total number of words per sentence) and uses TNWS to index the Sentence Data Array. SEARCH finds the next text word and places it in WORD [TNWS]. It then searches the dictionary for the text word. The exact form of the search will depend on the media of the dictionary.

If the word is in the dictionary file, all information will be added to the entry in Sentence Data Array.

If the word is not in the dictionary, the action taken will depend on the mode. If batch, the Sentence Data Array entry will be marked as containing an unknown word. The various items will be set to their default values:

NOMORE ← 1
NOPARTS ← 0
PART[*] ← all 0
NOSYLLABLES ← [# letters/3] + 1
NEGIND ← no
SYMBOLIND ← no
NOWORDS ← 1
NOREFS ← 1
EXAMPLE ← no
CLICHE ← no

If the run is in the interactive mode and the word is not in the dictionary, the user will be asked to supply the necessary information. If the user supplies the information, the supplied information will be used and saved in the dictionary for possible later use. If the user choosed not to supply the information, the various items will be assigned their default values as defined above.

INPUT: Next text word.

FILES ACCESSED: text file
dictionary file
cliches file
example file
input file (only if not in dictionary)

GLOBAL DATA: Updates TNWS, adds one entry to the Sentence Data Array. If word not in dictionary, updates NWNDB and NOTINDICT.

OUTPUT: SEARCH sets up one entry of the Sentence Data Array.

MODULES CALLED:

ERROR } if word not in dictionary and
SCANIN } mode is interactive

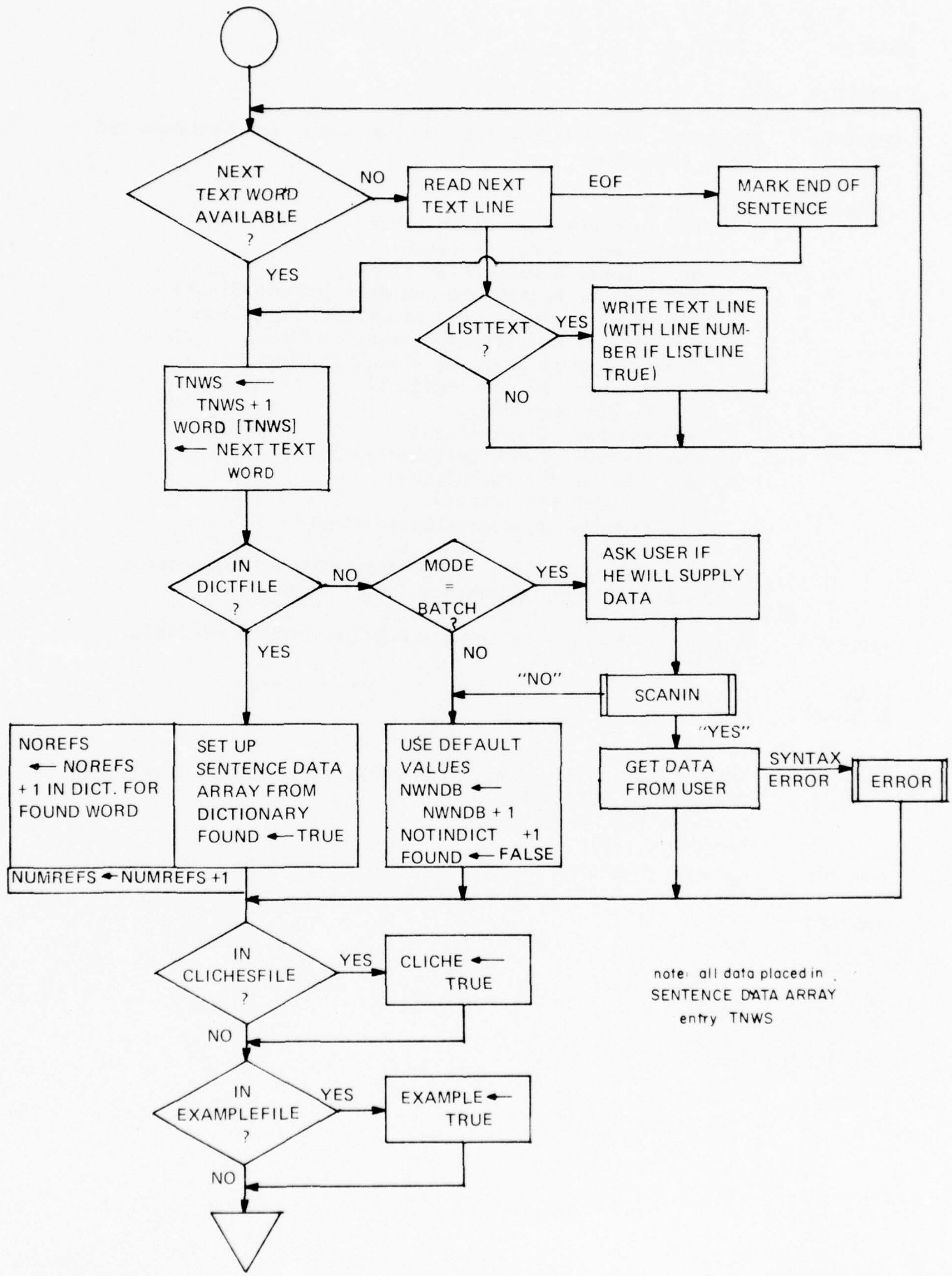
CALLING MODULES: None

COMMENT: The implementation difficulty of SEARCH is directly affected by:

1. dictionary form
2. complexity of "word-not-in-dictionary" code.

The entire problem of interactive dictionary update might be ignored for the first version. When it is added, the comments for the READ module apply.

If a new text line is required, SEARCH will read it and write the text if LISTTEXT is true.


NEXT TEXT WORD AVAILABLE ?
NO
YES
READ NEXT TEXT LINE
EOF
MARK END OF SENTENCE
LISTTEXT ?
YES
NO
WRITE TEXT LINE (WITH LINE NUMBER IF LISTLINE TRUE)
TNWS ← TNWS + 1
WORD [TNWS] ← NEXT TEXT WORD
IN DICTFILE ?
NO
YES
MODE = BATCH ?
YES
NO
ASK USER IF HE WILL SUPPLY DATA
SCANIN
"NO"
"YES"
NOREFS ← NOREFS + 1 IN DICT. FOR FOUND WORD
SET UP SENTENCE DATA ARRAY FROM DICTIONARY
FOUND ← TRUE
NUMREFS ← NUMREFS +1
USE DEFAULT VALUES
NWNDB ← NWNDB + 1
NOTINDICT +1
FOUND ← FALSE
GET DATA FROM USER
SYNTAX ERROR
ERROR
IN CLICHESFILE ?
YES
NO
CLICHE ← TRUE
note: all data placed in SENTENCE DATA ARRAY entry TNWS
IN EXAMPLEFILE ?
YES
NO
EXAMPLE ← TRUE

NAME: COUNT

NUMBER: 8

PURPOSE: The COUNT module maintains running counts for sentence and block summaries.

TECHNIQUE: COUNT updates

- NDWB (different words in block)
- TNWB (number words in block)
- TNMB (number morphemes in block)
- NAMB (number multi-word abbreviations in block)
- NASB (number single word abbreviations in block)
- NDNB (number of different nouns in block)
- TPSB (number of parts of speech in block)
- TNEB (number of elucidations)
- NORB (right branching)
- TSCB (number of syllables)
- OSWB (number of one syllable words)
- TNCB (number of characters)
- TCLB (number of clauses)
- NSWB (number of total abbreviations)

and adds the next entry in the BLOCKWORD table; updates BLOCKWORD table for words not in dictionary.

INPUT: Current Sentence Data Array entry (as defined by TNWS).

FILES ACCESSED: None

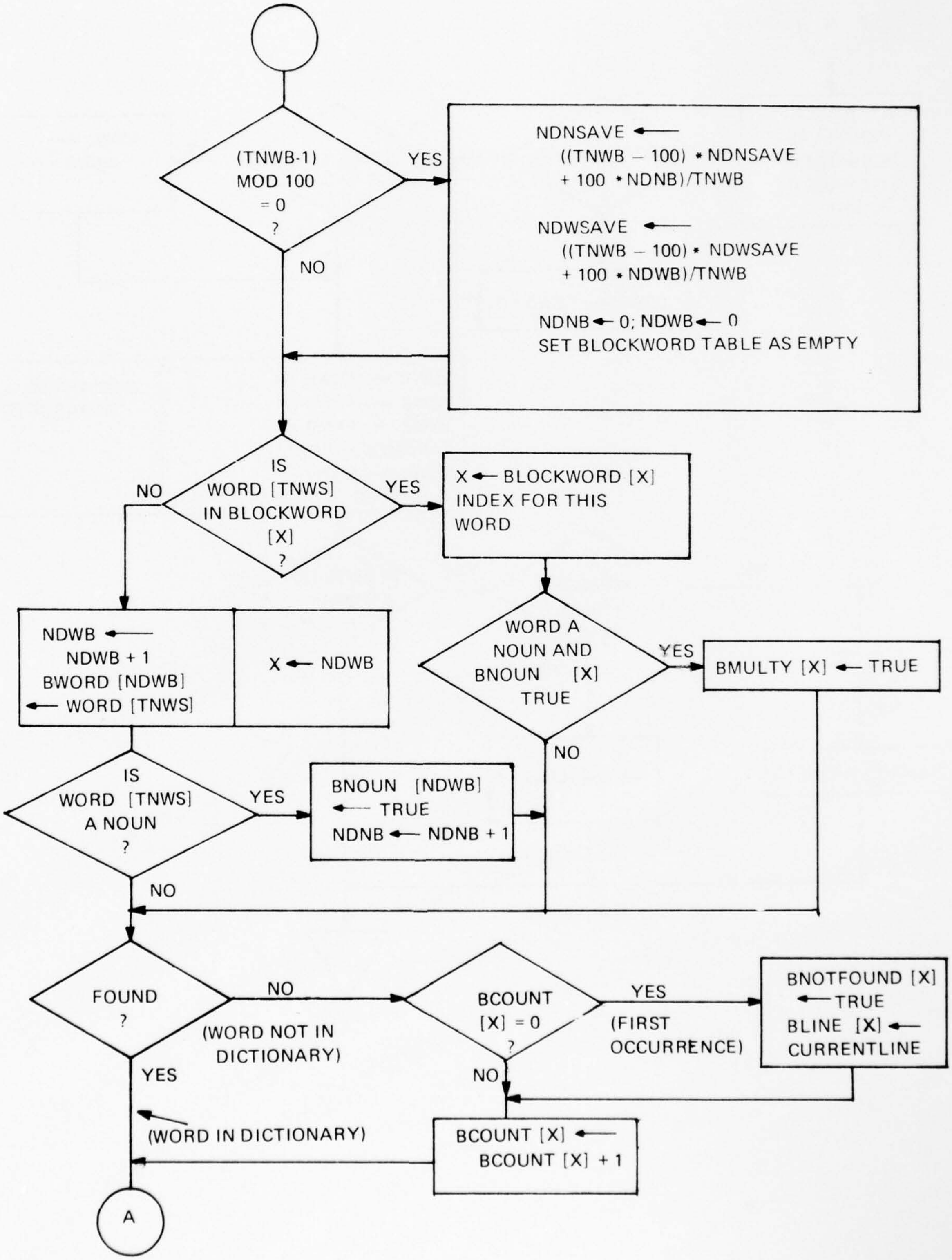
GLOBAL DATA: TNDWB, TNWB, TNMB, NAMB, NASB, NDNB, TPSB, TNEB, NORB, TCSB, OSWB, TNCB, TCLB, NSWB
Sentence Data Array, entry
Blockword table.

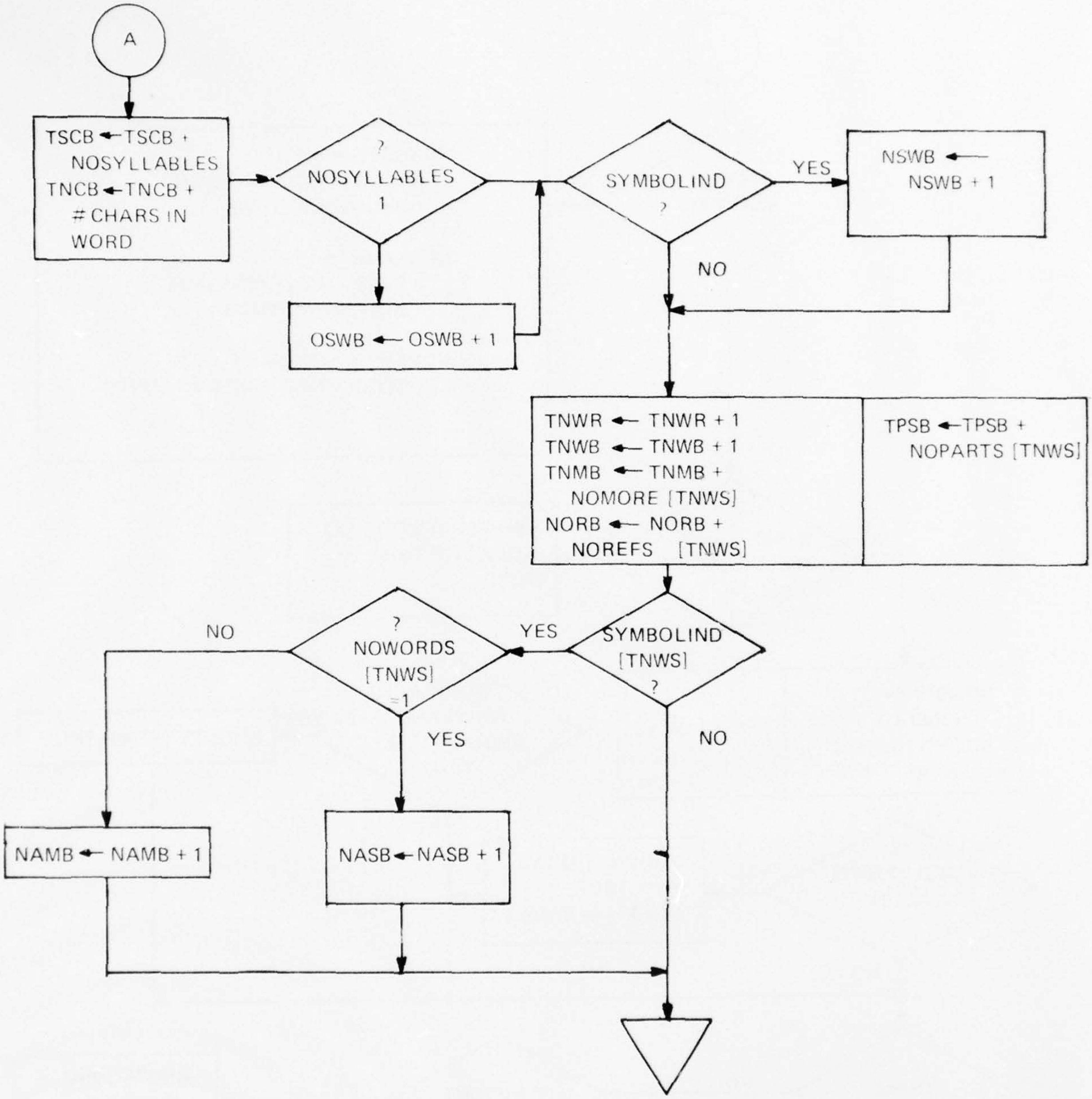
OUTPUT: Updated counters.

MODULES CALLED: None

CALLING MODULES: None

COMMENTS: None


(TNWB-1)
MOD 100
= 0
?
YES
NO
NDNSAVE ←
((TNWB – 100) * NDNSAVE
+ 100 * NDNB)/TNWB
NDWSAVE ←
((TNWB – 100) * NDWSAVE
+ 100 * NDWB)/TNWB
NDNB ← 0; NDWB ← 0
SET BLOCKWORD TABLE AS EMPTY
IS
WORD [TNWS]
IN BLOCKWORD
[X]
?
NO
YES
X ← BLOCKWORD [X]
INDEX FOR THIS
WORD
NDWB ←
NDWB + 1
BWORD [NDWB]
← WORD [TNWS]
X ← NDWB
WORD A
NOUN AND
BNOUN [X]
TRUE
YES
BMULTY [X] ← TRUE
NO
IS
WORD [TNWS]
A NOUN
?
YES
BNOUN [NDWB]
← TRUE
NDNB ← NDNB + 1
NO
FOUND
?
NO
(WORD NOT IN
DICTIONARY)
BCOUNT
[X] = 0
?
YES
(FIRST
OCCURRENCE)
BNOTFOUND [X]
← TRUE
BLINE [X] ←
CURRENTLINE
NO
YES
(WORD IN DICTIONARY)
BCOUNT [X] ←
BCOUNT [X] + 1
A


A
TSCB ← TSCB + NOSYLLABLES
TNCB ← TNCB + # CHARS IN WORD
? NOSYLLABLES 1
OSWB ← OSWB + 1
SYMBOLIND ?
YES
NSWB ← NSWB + 1
NO
TNWR ← TNWR + 1
TNWB ← TNWB + 1
TNMB ← TNMB + NOMORE [TNWS]
NORB ← NORB + NOREFS [TNWS]
TPSB ← TPSB + NOPARTS [TNWS]
NO
? NOWORDS [TNWS] =1
YES
SYMBOLIND [TNWS] ?
YES
NO
NAMB ← NAMB + 1
NASB ← NASB + 1

NAME: PARSE

NUMBER: 9

PURPOSE: The PARSE module is intended to parse fully each sentence presented to it. It attempts to reduce to a minimum the number of potential parse-trees and produces a representation of a parse-tree for each of the possible parses of the sentence.

TECHNIQUE: The PARSE module consists of a parser and a preparser. It makes every possible assignment of categories (parts of speech) to each word of the sentence. Each assignment is passed to the preparser, which either eliminates it or imposes on it a preliminary structure.

The parser, by means of repetitive applications of parsing rules (Appendix F) derives a single tree structure for each possible reading of the sentence. The tree structure is embodied in a PARSE-TABLE.

The parser maintains an array called SUBSTRING, whose elements are the categories at the top of the portion of the tree at any given state of completion. It attempts to match the categories at the left of SUBSTRING with the right sides of the parsing rules. When a match is found, the left side of the parsing rule is entered into the PARSE-TABLE, indicating that the tree has been extended upward.

When at a given point a match is impossible, a back-tracking routine deletes the last-entered category from PARSE-TABLE and resumes the attempt to find a match, starting with the rule immediately following the one (in the list of parsing rules) that was used in the entry of the newly-deleted rule.

A similar technique is used in order to find additional parses of the same sentence using the same sequence of categories. A copy of PARSE-TABLE, which is complete for a given parse is made, and deletions and additions to this copy are made to get additional parses. The process is repeated until no additional parses are possible, or until the maximum number of parses (see PARSE LIMIT SPEC, Appendix C).

GET-SUBSTRING finds the arrays SUBSTRING and SUBSTRING-LINES. If FIRST-LINE is not 1, then only the categories from the left-most point are desired, so they cab be erased in backtracking.

FOLLOW-PATH starts with the word I of the current sentence, and goes upward from there through the PARSE-TABLE. SIGNAL is set to 1 if this is the left-most path to its highest point, and to -1 otherwise. CAT and LINE are the category at this highest point and the line is PARSE-TABLE on which it is entered.

BEGIN-RULES tries to find a match for the string SUBSTRING. It begins with the rule numbered FIRST-RULE and goes through each rule from that point on. If a rule has a length LENGTH(I), then the first LENGTH(I) symbols from SUBSTRING are used. The procedure calls ENTER when a match is found, and BAD-MATCH when no match is found.

ENTER begins a new line in PARSE-TABLE and places pointers on previous lines which correspond to categories which lead upward to the new category. It also places a 1 in column 3 of the line of the left-most of these categories, and -1 in column 3 of the others. On the new line, it enters the category in column 1, 0 in columns 2 and 3, and the number of the rule just used in column 4.

GOOD-MATCH checks to see if the tree has been completed without arriving at S. If this happens, BACKTRACK is called. If the tree is completed with S, then, FINISHED-PARSE is called. Otherwise, parsing is resumed by getting a new SUBSTRING and calling BEGIN-RULES.

FINISHED-PARSE is called each time a parse has been found.

BAD-MATCH drops the first element of SUBSTRING, if this is possible, and calls BEGIN-RULES to look for new matches. if SUBSTRING has only 1 element, BACKTRACK is called.

BACKTRACT erases all references in PARSE-TABLE to its last line, and erases the last line itself. It then prepares to look for a new match at the point where the last one was found, but starting with the next rule in the list. If there is noting to erase, NO-PARSE is called.

MAIN is the main procedure.

NO-PARSE is called when no parse can be found for the given assignments of categories to the words of the sentence.

INPUT: The output is the Sentence Data Array from the SEARCH module.

FILES ACCESSED: SUBSTRING - an array whose elements are the categories at the top of the current sub-tree

SUBSTRING-LINES - An array whose elements are the numbers of the lines of PARSE-TABLE which correspond to the elements of SUBSTRING

SUBSTRING-LENGTH - the number of elements in SUBSTRING

NUMWDS - the number of words in the current sentence

NUMRULES - the number of context-free rules

RULE - is a two dimensional array which has a row for each rule, and each row has as many elements as there are categories in that rule. For example, the row corresponding to S ← NP VP has two elements.

PARSE-TABLE - is a two dimensional array which is a record of the developing parse

PARSE-LINE - is the line of the PARSE-TABLE in current use

LENGTH - an array whose ith element is the number of categories in the ith rule.

LEFT - an array whose ith element is the category to the left of the arrow in the ith rule

NUMASS - the number of possible combinations of assignments of categories to the words of the sentence.

OUTPUT: The PARSE module produces a copy of the PARSE-TABLE array for each possible parse of the sentence and, if possible, a sentence diagram.

MODULES CALLED: None

CALLING MODULES: None

PREPARSER

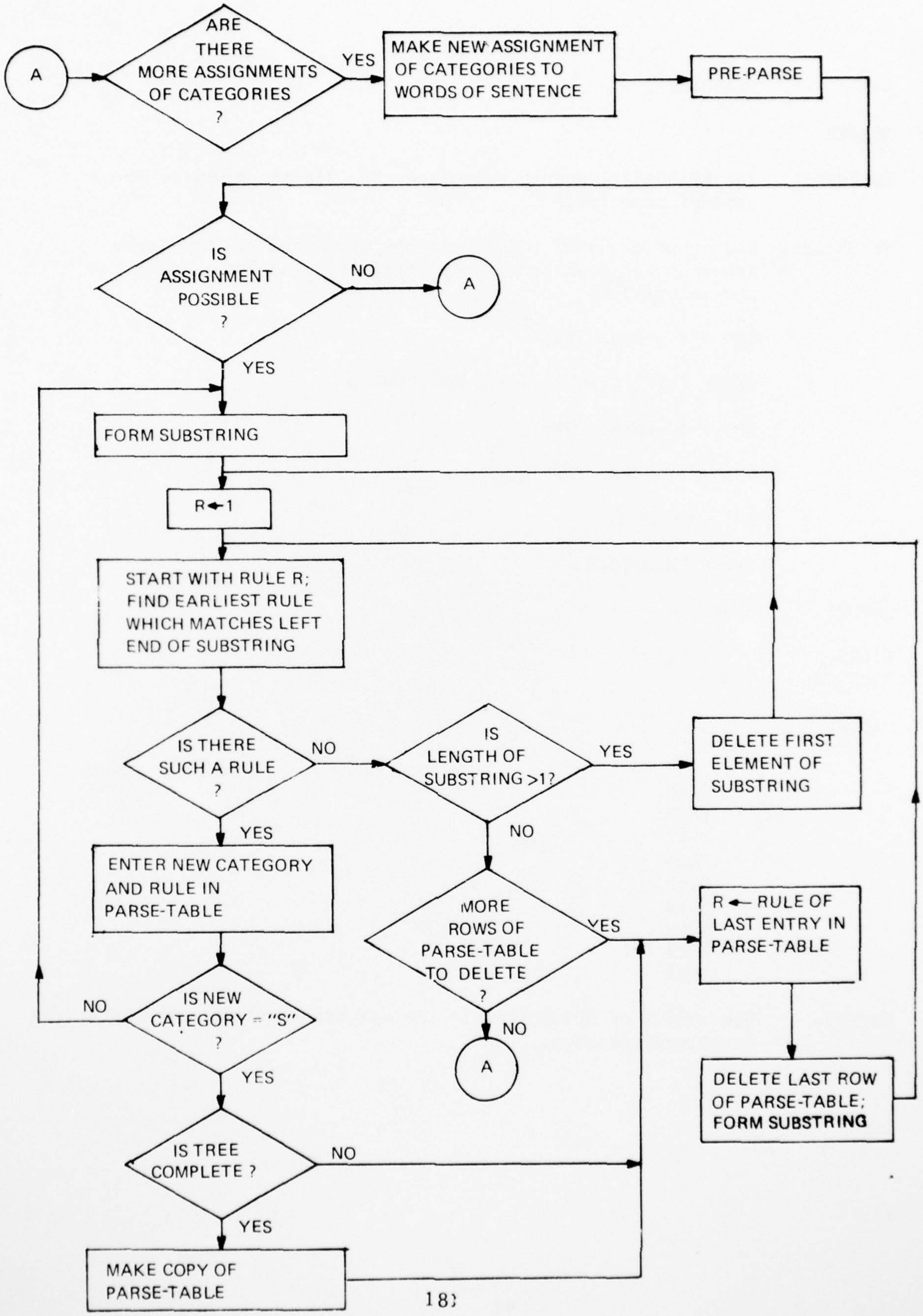
The preparser has two goals:

1. To eliminate certain impossible sequences of categories entry into the parser.

2. To apply a limited number of frequently applicable context-sensitive parsing rules, thereby giving a preliminary structure to the sentence. This preliminary structure is not subsequently altered by the parser.

In order to accomplish the first goal, every member of a list of impossible categories is compared whenever a new sequence of categories comes under consideration. If at any point there is a match, the entire current sequence is eliminated and the next assignment of categories to the words of the sentence is made.

To accomplish the second goal, the list of parsing rules is compared in turn with the current sequence of categories. When a match is found, the new left category is entered into PARSE-TABLE in the same manner as this entry is performed by the parser.


A
ARE THERE MORE ASSIGNMENTS OF CATEGORIES ?
YES
MAKE NEW ASSIGNMENT OF CATEGORIES TO WORDS OF SENTENCE
PRE-PARSE
IS ASSIGNMENT POSSIBLE ?
NO
A
YES
FORM SUBSTRING
R ← 1
START WITH RULE R; FIND EARLIEST RULE WHICH MATCHES LEFT END OF SUBSTRING
IS THERE SUCH A RULE ?
NO
IS LENGTH OF SUBSTRING >1?
YES
DELETE FIRST ELEMENT OF SUBSTRING
YES
NO
ENTER NEW CATEGORY AND RULE IN PARSE-TABLE
MORE ROWS OF PARSE-TABLE TO DELETE ?
YES
R ← RULE OF LAST ENTRY IN PARSE-TABLE
NO
IS NEW CATEGORY = "S" ?
NO
A
YES
DELETE LAST ROW OF PARSE-TABLE; FORM SUBSTRING
IS TREE COMPLETE ?
NO
YES
MAKE COPY OF PARSE-TABLE

NAME: MEASURE/SI

NUMBER: 10

PURPOSE: The MEASURE/SI module calculates the six structure of intellect measures.

TECHNIQUE: All input required to compute the structure of intellect measures has been previously obtained. The actual measures are calculated:

CMU ← 1 - NDWB/TNWB

CMR ← (NSNB/(TNSB- 1) + NORB)/TNWB

MMU ← 1 -(NDNB/TNWB)

ESI ← 1 -(NSWB/TNWB)

NMI ←(TNWB/TPSB)

DMU ← TNEB/TNSB

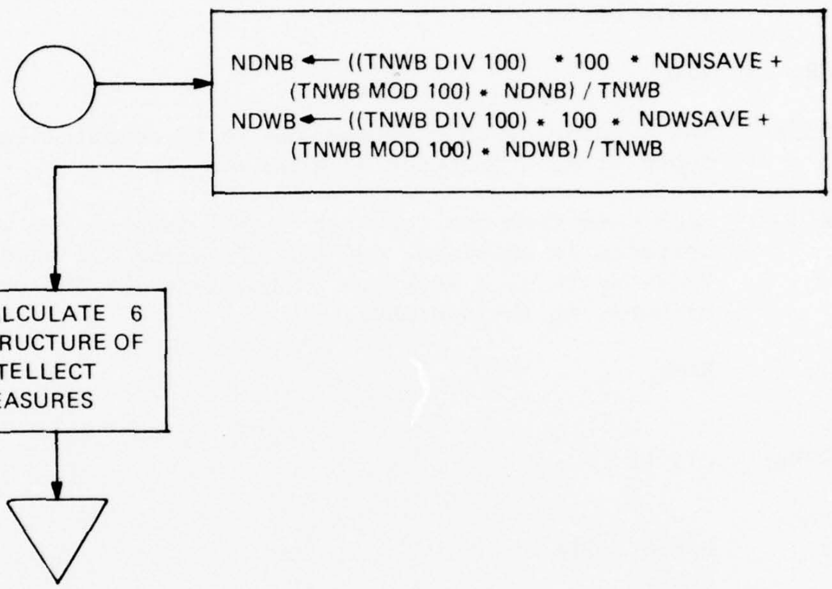
INPUT: None

FILES ACCESED: None

GLOBAL DATA: Uses:

NDWB
TNWB
NSNB
TNSB
NQRB
NDNB
NSWB
NPPB
TNEB

OUTPUT: The output of MEASURE/SI is the six nonnormalized structure of intellect measures.


NDNB ← ((TNWB DIV 100) * 100 * NDNSAVE +
(TNWB MOD 100) * NDNB) / TNWB
NDWB ← ((TNWB DIV 100) * 100 * NDWSAVE +
(TNWB MOD 100) * NDWB) / TNWB
CALCULATE 6 STRUCTURE OF INTELLECT MEASURES

NAME: MEASURE/P

NUMBER: 11

PURPOSE: The MEASURE/P module calculates the seven psycholinguistic comprehensibility measures. The submodules which each compute measure are individually described as numbers 11A - 11G below:

NAME: YNGVE DEPTH (YD)

NUMBER: 11A

PURPOSE: The purpose of this subroutine is to compute the Yngve depth of each sentence in a block.

TECHNIQUE: Each path from the top-most symbol down to the words in the sentence is examined, and the YDs along all paths summed. Yngve depth of a sentence is the same divided by the number of words in the sentence.

INPUT: None

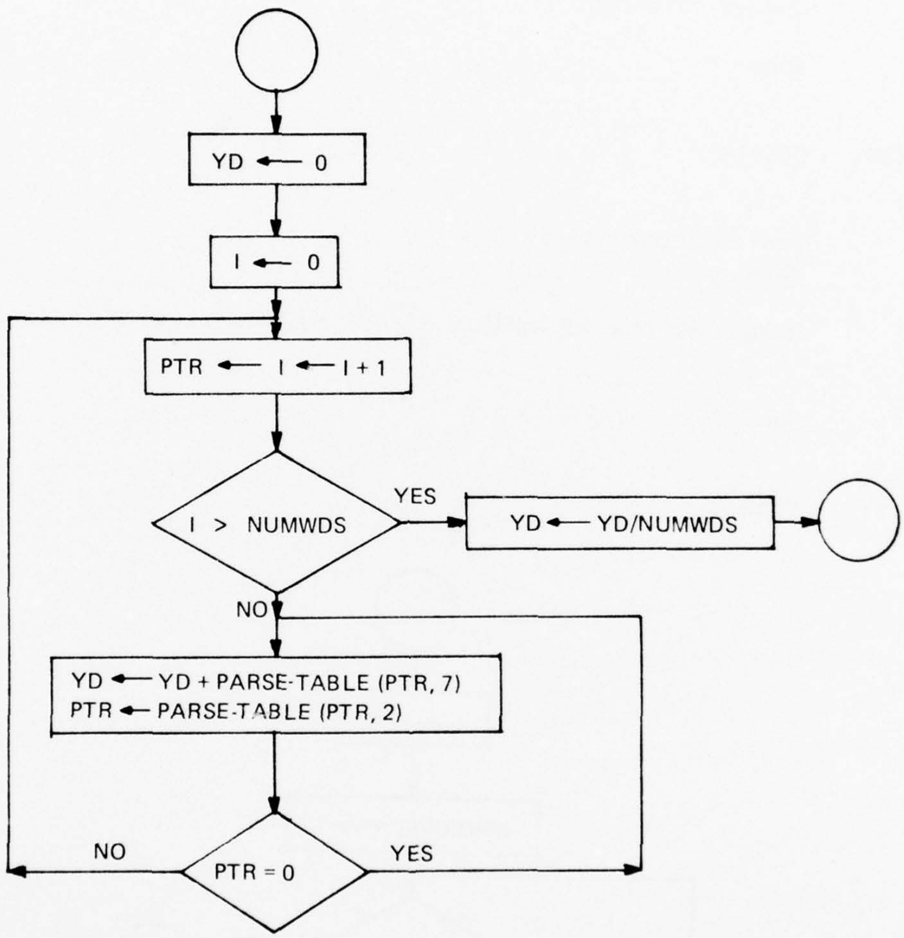
FILES ACCESSED: Cliché

GLOBAL DATA: PARSE TABLE
NUMWDS

OUTPUT: Output of the sub-module is the nonnormalized YD value.

MODULES CALLED: None

CALLING MODULES: None


YD ← 0
I ← 0
PTR ← I ← I + 1
YES
I > NUMWDS
YD ← YD/NUMWDS
NO
YD ← YD + PARSE-TABLE (PTR, 7)
PTR ← PARSE-TABLE (PTR, 2)
NO
PTR = 0
YES

NAME: MORPHEME DEPTH (MD)

NUMBER: 11B

PURPOSE: Calculation of the psycholinguistic measure MD.

TECHNIQUE: The number of morphemes in the sentence is divided by the number of words.

INPUT: None

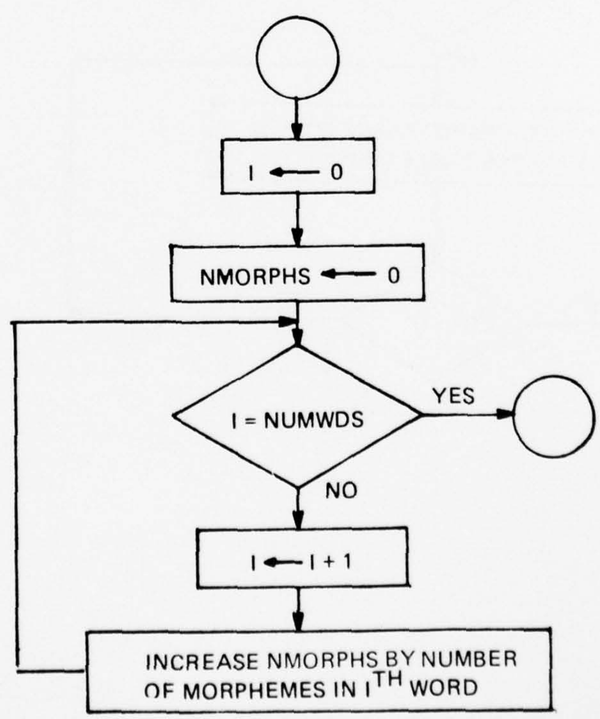
FILES ACCESSED: Cliche

GLOBAL DATA: Sentence Data Array
NUMWDS

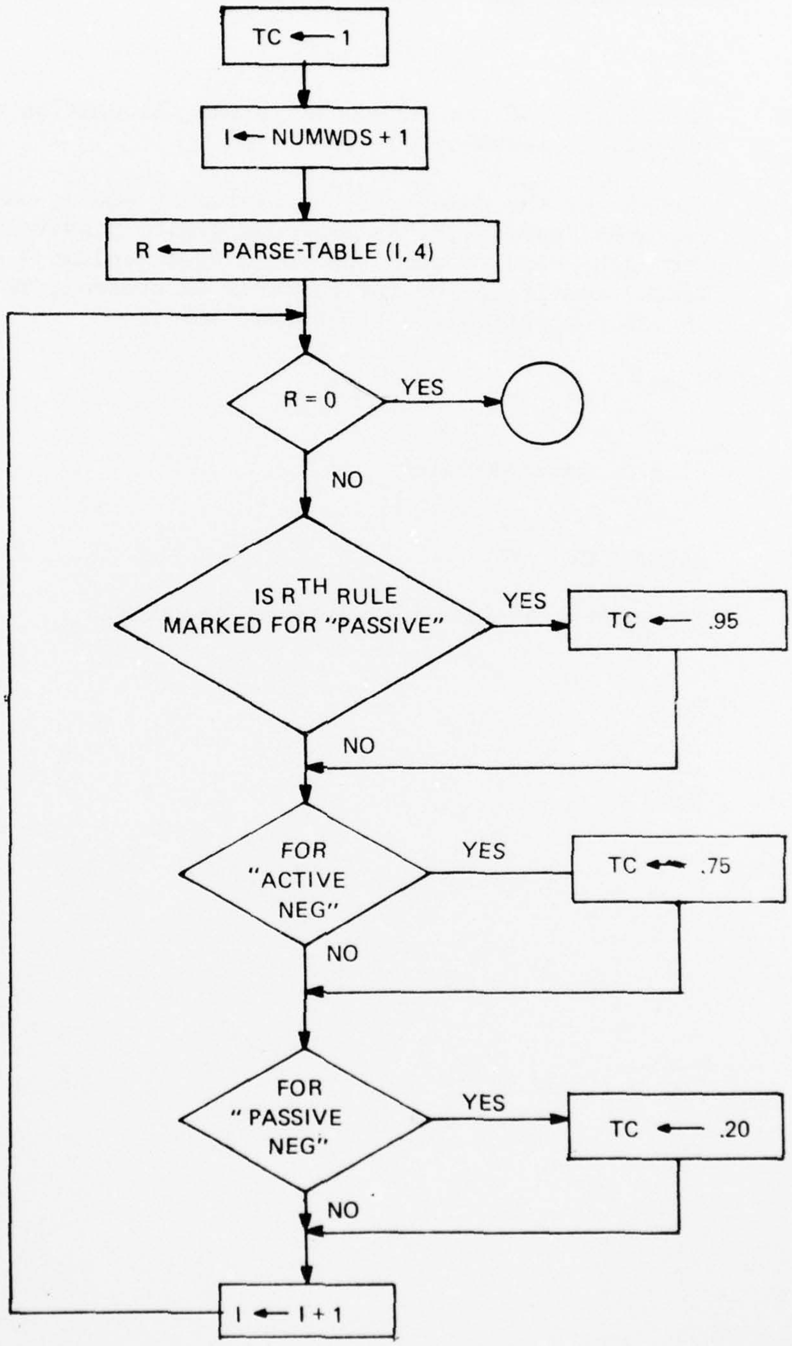
OUTPUT: Output of the sub-module is the nonnormalized MD value.

MODULES CALLED: None

CALLING MODULES: None



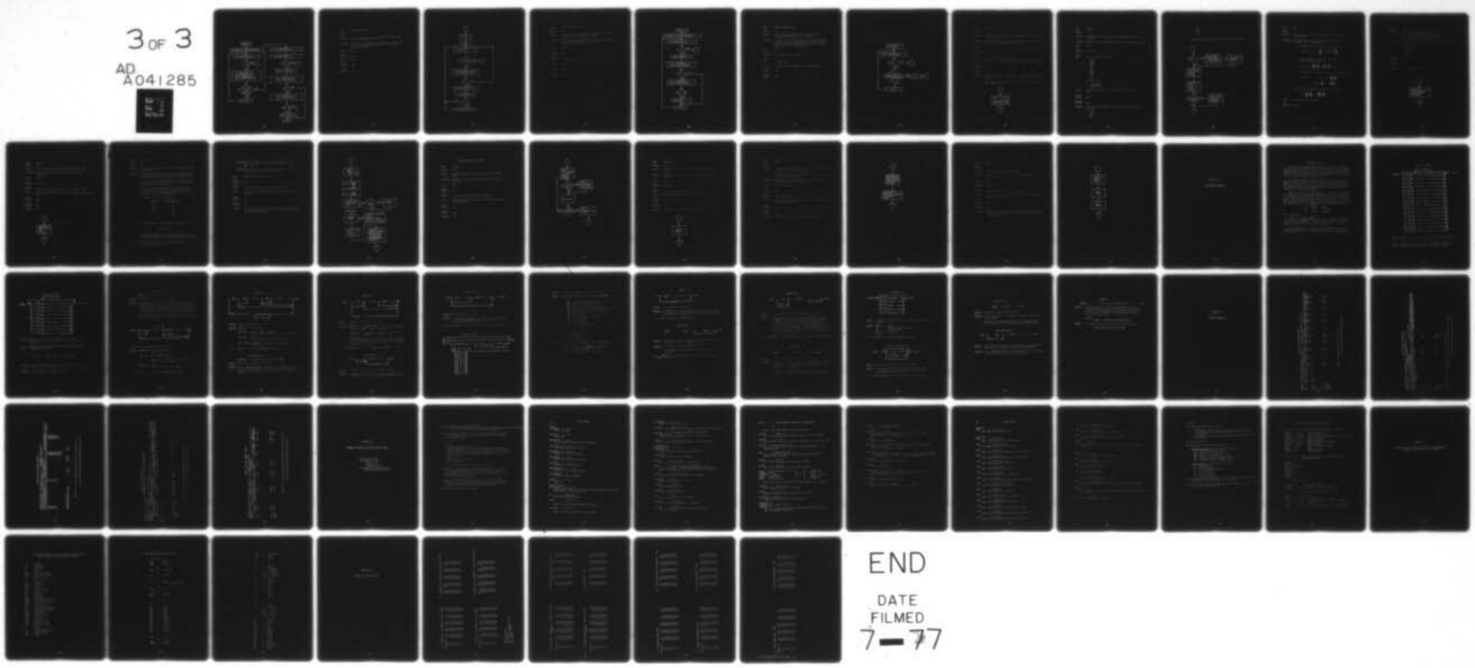
NAME:	TRANSFORMATIONAL COMPLEXITY (TC)
NUMBER:	11C
PURPOSE:	The purpose of the submodule is the calculation of the psycholinguistic measure TC.
TECHNIQUE:	Certain of the rules will be listed as having one of the features "passive," "active-negative," "passive-negative." For each parse, search the rules used (column 4 of PARSE-TABLE) and if any of the features is present, TC is .95, .75, or .20 respectively. Otherwise, TC= 1.
INPUT:	None
FILES ACCESSED:	File of parsing rules
GLOBAL DATA:	PARSE TABLE
OUTPUT:	The output is the nonnormalized TC value.
MODULES CALLED:	None
CALLING MODULES:	None


TC ← 1
I ← NUMWDS + 1
R ← PARSE-TABLE (I, 4)
R = 0
YES
NO
IS R^TH RULE MARKED FOR "PASSIVE"
YES
TC ← .95
NO
FOR "ACTIVE NEG"
YES
TC ← .75
NO
FOR "PASSIVE NEG"
YES
TC ← .20
NO
I ← I + 1

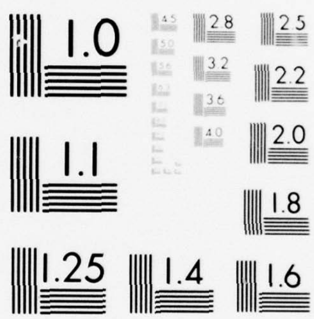
NAME:	CENTER EMBEDDING (CE)
NUMBER:	11D
PURPOSE:	The purpose of this submodule is to calculate the value of NNPS, required for the subsequent determination of CE.
TECHNIQUE:	The subject noun phrase of the sentence is located, and the number of phrases branching off to the right from this phrase is determined.
INPUT:	None
FILES ACCESSED:	None
GLOBAL DATA:	PARSE TABLE NUMWDS
OUTPUT:	NNPS
MODULES CALLED:	None
CALLING MODULES:	None

AD-A041 285 APPLIED PSYCHOLOGICAL SERVICES INC WAYNE PA F/G 5/7
STUDIES AND DESIGN SPECIFICATIONS FOR COMPUTERIZED MEASUREMENT --ETC(U)
OCT 76 A I SIEGEL, A R WILLIAMS, W J LAPINSKY F41609-75-C-0037
UNCLASSIFIED AFHRL-TR-76-77 NL

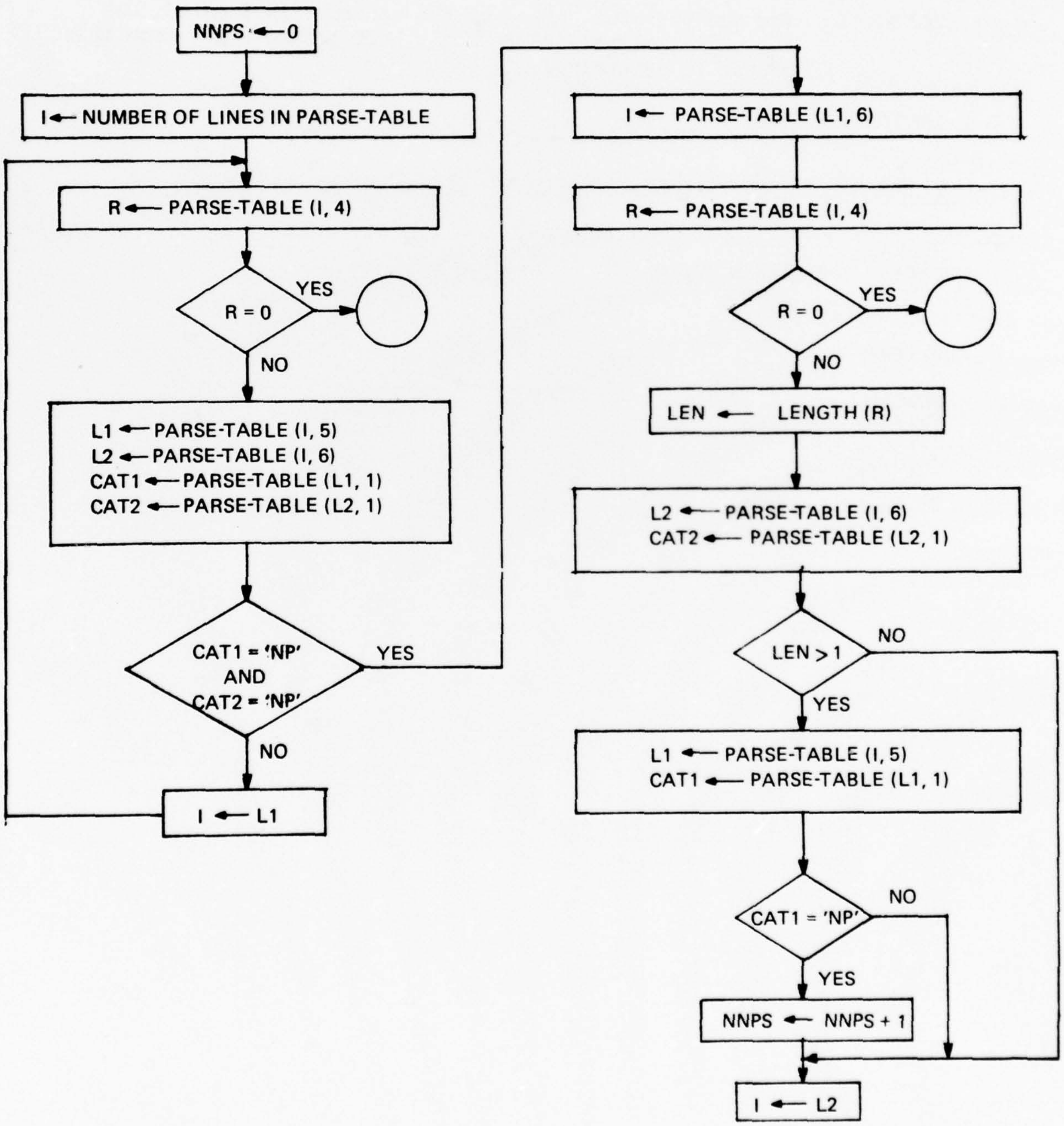
3 OF 3
AD
A041285



END
DATE
FILMED
7-77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


NNPS ← 0
I ← NUMBER OF LINES IN PARSE-TABLE
R ← PARSE-TABLE (I, 4)
R = 0
YES
NO
L1 ← PARSE-TABLE (I, 5)
L2 ← PARSE-TABLE (I, 6)
CAT1 ← PARSE-TABLE (L1, 1)
CAT2 ← PARSE-TABLE (L2, 1)
CAT1 = 'NP' AND CAT2 = 'NP'
YES
NO
I ← L1
I ← PARSE-TABLE (L1, 6)
R ← PARSE-TABLE (I, 4)
R = 0
YES
NO
LEN ← LENGTH (R)
L2 ← PARSE-TABLE (I, 6)
CAT2 ← PARSE-TABLE (L2, 1)
LEN > 1
NO
YES
L1 ← PARSE-TABLE (I, 5)
CAT1 ← PARSE-TABLE (L1, 1)
CAT1 = 'NP'
NO
YES
NNPS ← NNPS + 1
I ← L2

NAME: LEFT BRANCHING (LB)

NUMBER: 11E

PURPOSE: The submodules purpose is to compute the value of NCLS, required for computation of LB.

TECHNIQUE: The subject noun phrase of the sentence is located, and the number of phrases branching off to the left from this phrase is determined.

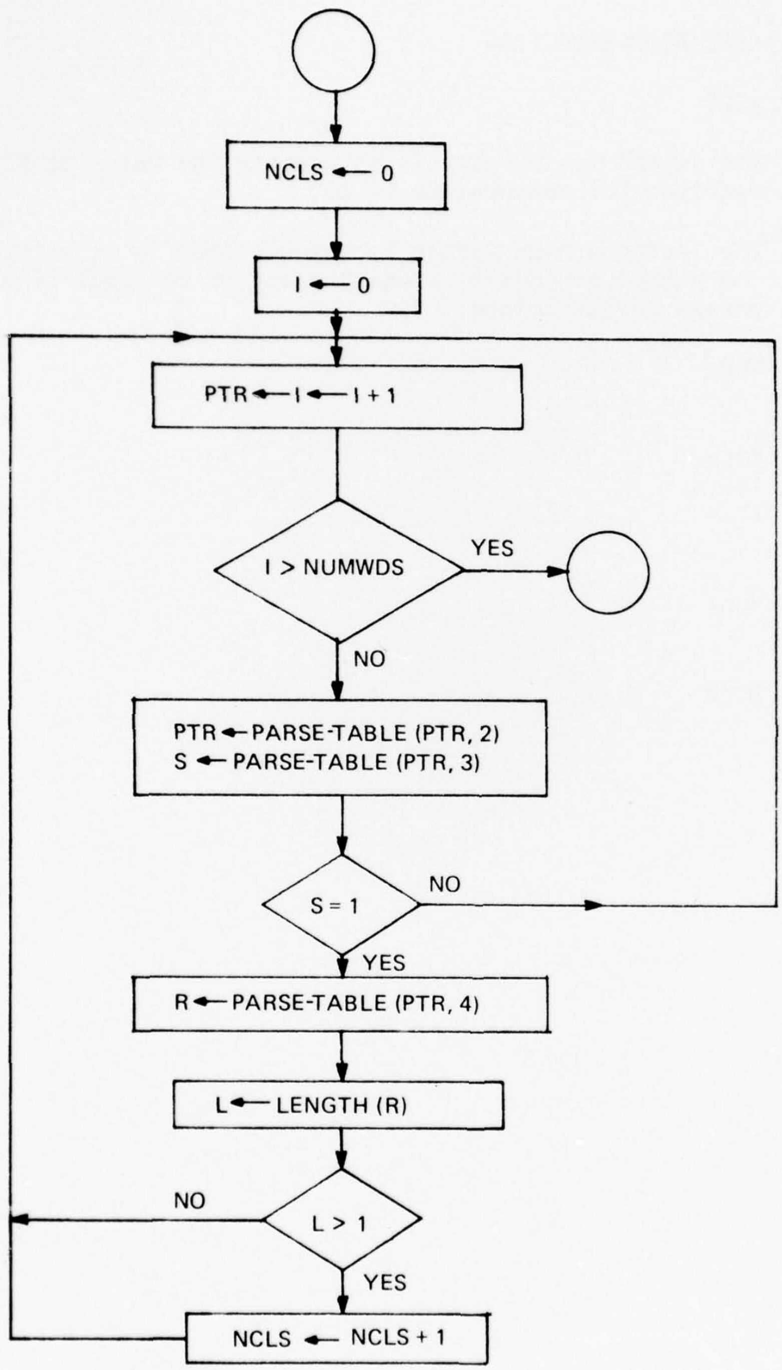
INPUT: None

FILES ACCESSED: None

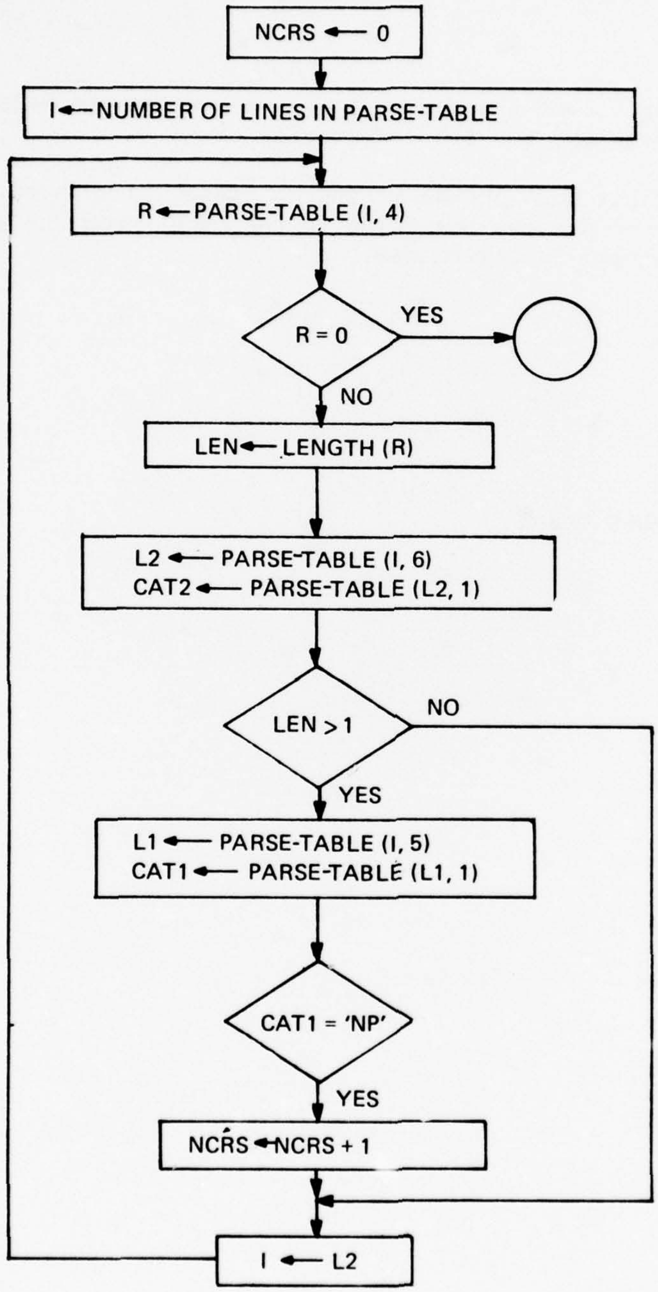
OUTPUT: NCLS

MODULES CALLED: None

CALLING MODULES: None


NCLS ← 0
I ← 0
PTR ← I ← I + 1
I > NUMWDS
YES
NO
PTR ← PARSE-TABLE (PTR, 2)
S ← PARSE-TABLE (PTR, 3)
S = 1
NO
YES
R ← PARSE-TABLE (PTR, 4)
L ← LENGTH (R)
L > 1
NO
YES
NCLS ← NCLS + 1

NAME:	RIGHT BRANCHING (RB)
NUMBER:	11F
PURPOSE:	The purpose of the submodule is to compute the value of NCRS required for the computation of RB.
TECHNIQUE:	The object noun phrase of the sentence is located. The number of phrases branching to the right from the object noun phrase is determined.
INPUT:	None
FILES ACCESSED:	None
GLOBAL DATA:	Uses: PARSE TABLE
OUTPUT:	NCRS
MODULES CALLED:	None
CALLING MODULES:	None



NCRS ← 0
I ← NUMBER OF LINES IN PARSE-TABLE
R ← PARSE-TABLE (I, 4)
R = 0
YES
NO
LEN ← LENGTH (R)
L2 ← PARSE-TABLE (I, 6)
CAT2 ← PARSE-TABLE (L2, 1)
LEN > 1
NO
YES
L1 ← PARSE-TABLE (I, 5)
CAT1 ← PARSE-TABLE (L1, 1)
CAT1 = 'NP'
YES
NCRS ← NCRS + 1
I ← L2

NAME: DELETED COMPLEMENT (DC)

NUMBER: 11G

PURPOSE: The purpose of the submodule is calculation of the psycholinguistic measure DC of a sentence.

TECHNIQUE: Certain of the rules will be listed as having a deleted complement feature. For each parse, the rules are searched (column 4 of PARSE TABLE) and if any rule used has this feature then DC= 0. Otherwise, DC= 1.

INPUT: None

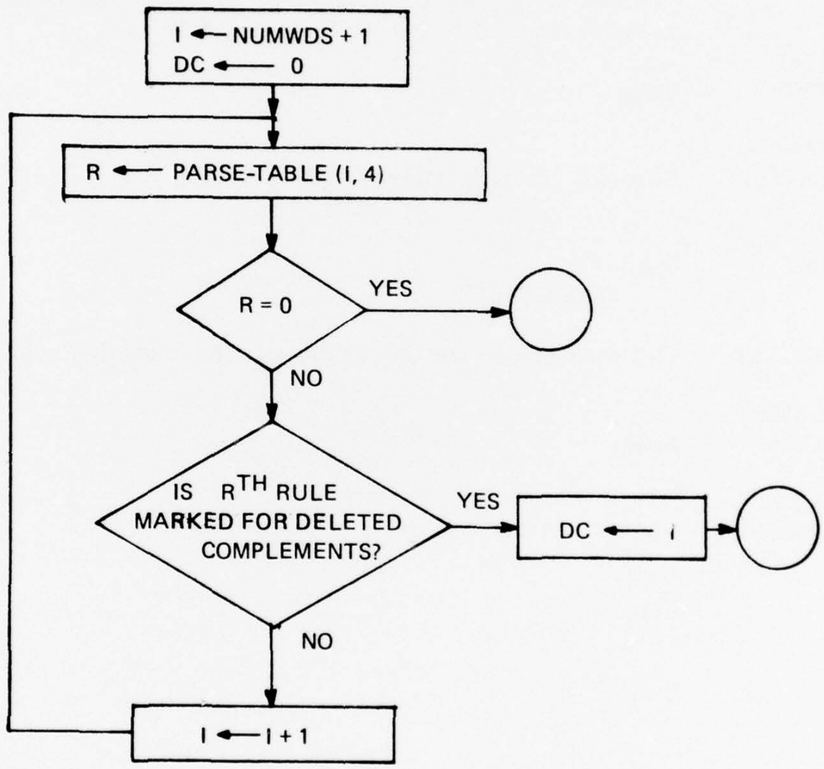
FILES ACCESSED: File of parsing rules: RULE

GLOBAL DATA: Uses:
PARSE TABLE

OUTPUT: The output is the DC value of the examined sentence.

MODULES CALLED: None

CALLING MODULES: None


I ← NUMWDS + 1
DC ← 0
R ← PARSE-TABLE (I, 4)
R = 0
YES
NO
IS R^TH RULE MARKED FOR DELETED COMPLEMENTS?
YES
DC ← I
NO
I ← I + 1

NAME: SENTSUM

NUMBER: 12

PURPOSE: The SENTSUM module cummulates and summarizes the results of a single sentence.

TECHNIQUE: SENTSUM updates all block level statistics for each sentence that has not been updated previously.

INPUT: None

FILES ACCESSED: None

GLOBAL DATA: TNSB updated, NCRB, NCLB, NCRS, NCLS.

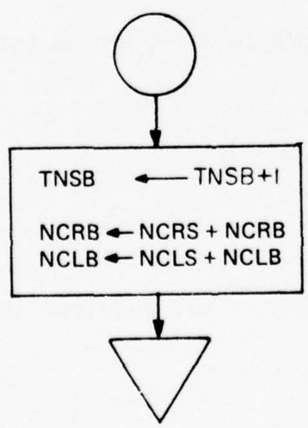
OUTPUT: The output of the SENTSUM module is the updated global items.

MODULES CALLED: None

CALLING MODULES: None

COMMENT: The work of SENTSUM is distributed in other modules, primarily COUNT.

NOTE: NSNB is calculated at the end of a block, not at the end of each sentence.



NAME: SENTOUT

NUMBER: 13

PURPOSE: The SENTOUT lists the result of processing each sentence in a block.

TECHNIQUE: SENTOUT is a report generator using values previously calculated.

INPUT: None

FILES ACCESSED: OUTPUT FILE

GLOBAL DATA: Uses: (for each sentence)

TNSB
TNWS
NPPS
NESS
YDS
TCS
NNPS
NCRS
NCLS
DCS

For header line:

BLOCKNUMBER
TEXTFILE name

OUTPUT: The output from SENTOUT is a report on each sentence in a block.

MODULES CALLED: None

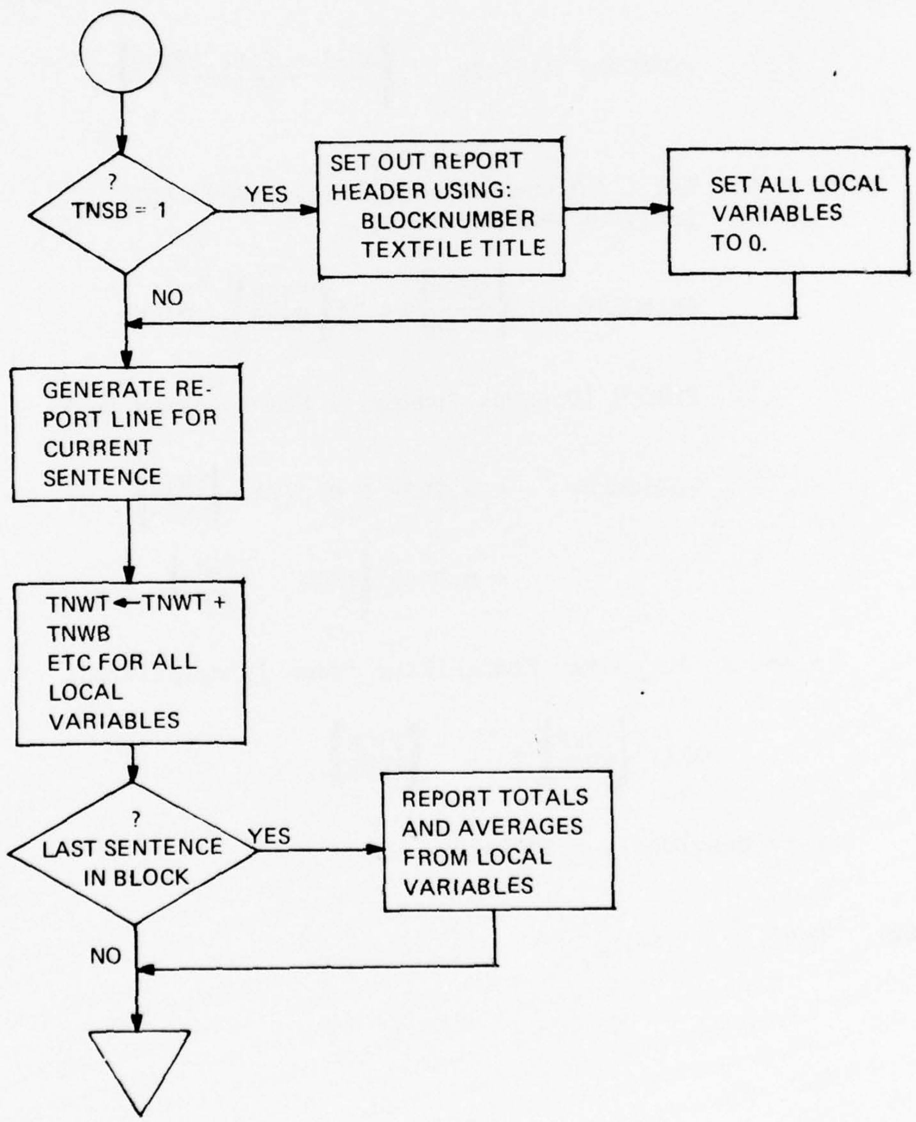
CALLING MODULES: None

COMMENTS: Items used only by SENTOUT but required to be saved between calls are:

NPST
NEST
YDT
TCT

NNPT
NCRT
NCLT
DCT

Each item is the total of the corresponding item with the final "T" replaced by "S."



NAME: RGL

NUMBER: 14

PURPOSE: The RGL module calculates the reading grade level for a block of text.

TECHNIQUE: The reading level is calculated by three different formulas. The selected formulas are:

1. FORECAST [Caylor, Sticht, Fox, & Ford; 1972]

$$\text{FORECAST RGL} = 20 - \left(\frac{\text{OSWB} \times (150/\text{TNWB})}{10}\right)$$

2. ARI (Automatic Readability Index) RGL [SMITH & SENTER; 1966]

$$\text{ARIRGL} = 0.5 \times \left(\frac{\text{TNWB}}{\text{TNSB}}\right) + 4.71 \left(\frac{\text{TNCB}}{\text{TNWB}}\right) - 21.43$$

3. FLESCH [Powers, Sumner, & Kearl (based on Flesch); 1958]

$$\text{FLESCHRL} = -2.2029 + 0.0778 \left(\frac{\text{TNWB}}{\text{TNSB}}\right) + 0.0455 \left(\text{TSCB} \cdot \frac{100}{\text{TNWB}}\right)$$

Also, an Automated Readability Index is calculated:

$$\text{ARI} = \left(\frac{\text{TNWB}}{\text{TNSB}}\right) + 9 \left(\frac{\text{TNCB}}{\text{TNWB}}\right)$$

INPUT: Data previously calculated.

FILES ACCESSED: None

GLOBAL
DATA: TNWB total number words in block
OSWB number of one syllable words in block
TNSB total number of sentences in block
TNCB total number of characters (letters) in block
TSCB total number of syllables in block

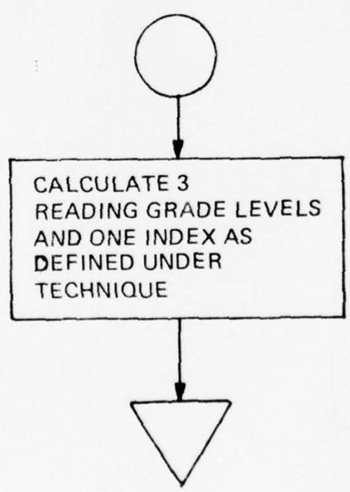
OUTPUT: The output of the RGL module is two reading grade levels and two indexes:

FORCASTRGL
ARIRGL
FLESCHRL
ARI

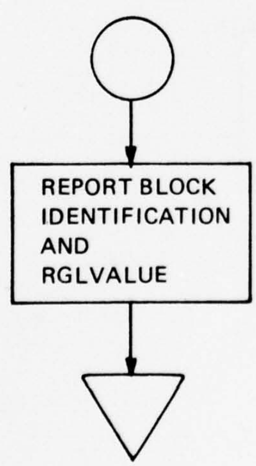
MODULES
CALLED: None

CALLING
MODULES: None

COMMENTS: The RGL and RGLOUT modules can be combined.



NAME:	RGLOUT
NUMBER:	15
PURPOSE:	RGLOUT displays the reading grade level for each block.
TECHNIQUE:	RGLOUT is a report generator using values previously calculated.
INPUT:	None
FILES ACCESSED:	OUTPUT FILE
GLOBAL DATA:	Uses three RGL values and indexes as calculated by RGL.
OUTPUT:	The output of RGLOUT is a report on the block's reading grade level.
MODULES CALLED:	None
CALLING MODULES:	None
COMMENTS:	None



NAME: BLOCKSUM

NUMBER 16

PURPOSE: The BLOCKSUM module performs the final calculation of block level values.

TECHNIQUE: BLOCKSUM calculates NSNB and normalizes all measures previously calculated. The measures are normalized by table interpolation. The tables are provided in Appendix F.

For values producing a normalized measure less than 5%, 5% will be used. For values producing a normalized measure greater than 95%, 95% will be used.

For values where the supplied interpolation tables reach the extremes of 0 and/or 1 over a range of values, the value of the corresponding measure must be set equal to the center of the band of extreme values.

For example, with the following table:

Measure	Normalized Value
1.0	90
1.0	80
1.0	70
1.0	60
.8	50
.6	40
.	.
.	.
	.

A measure value of 1.0 would yield a normalized value of

$$75 = \frac{90 - 60}{2}$$

The composite index formula depends on the settings of READERHIGH, READERLOW, and READERHILOW. One or more may be true, and for each one true a composite index is calculated:

COMPINDEXHIGH= - .132 * MMUN + .171 * ESIN + .418 * YDN + .397 * TCN + .302 * CMUN + .089 * MDN - . 320 * DMUN + .167 * RBN - .509

```
COMPINDEXHILO= .132 * MMUN + .164 * ESIN + .200 * YDN - .207 *
        SEN + .250 * RBN - .151 * MDN - .289 * NMIN - .074 *
        LBN - .003

COMPINDEXLOW= .169 *CMUN - .173 *MMUN + .156 *ESIN + .190 *DMUN
        + .335 *YDN + .260 *TCN + .242 *RBN - .344
```

INPUT: Results from previous modules

FILES ACCESSED: None

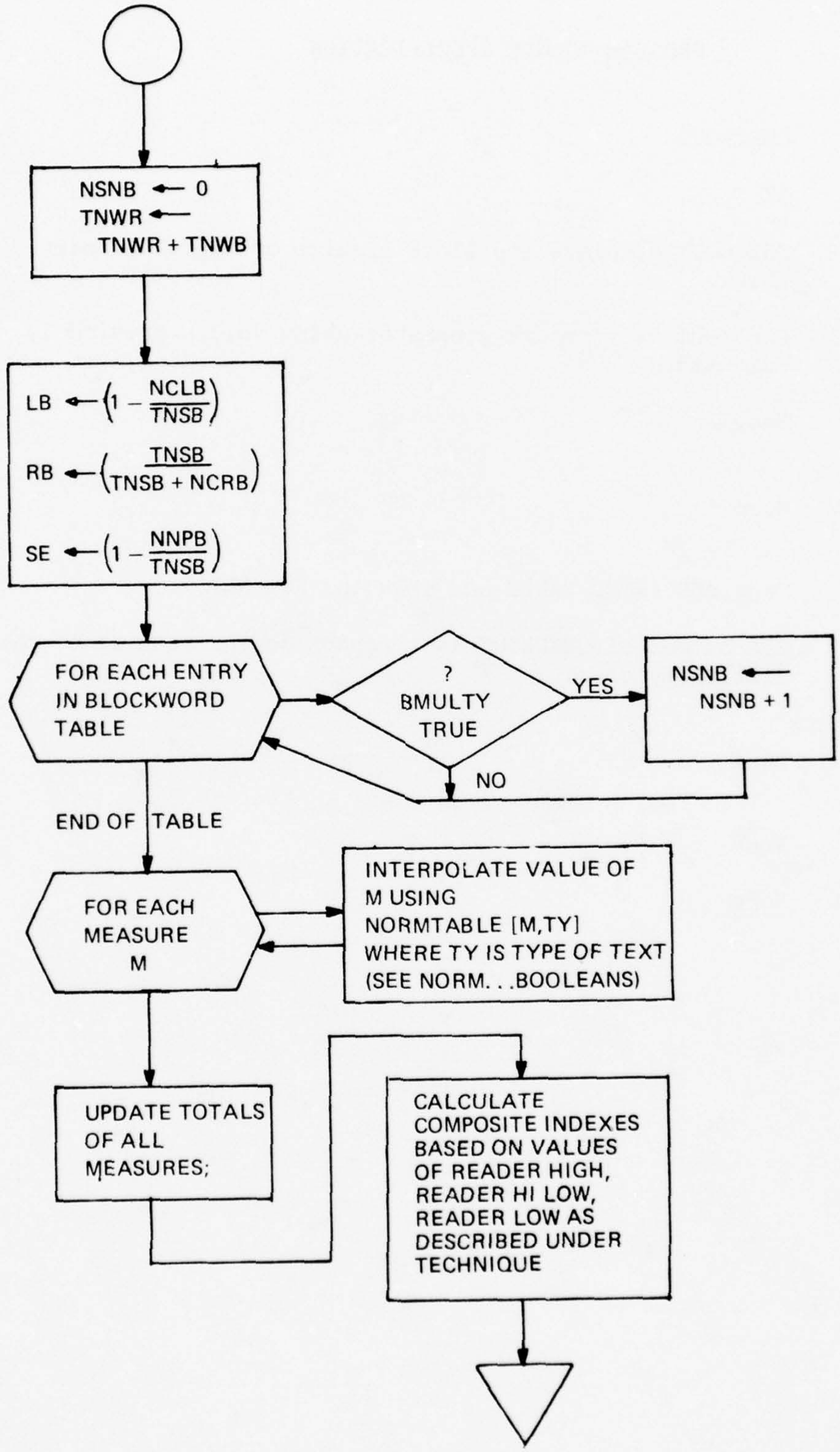
GLOBAL DATA: Updates NSNB and calculates normalized measure.

OUTPUT: The output of the BLOCKSUM module is the normalized measures and NSNB.

MODULES' CALLED: None

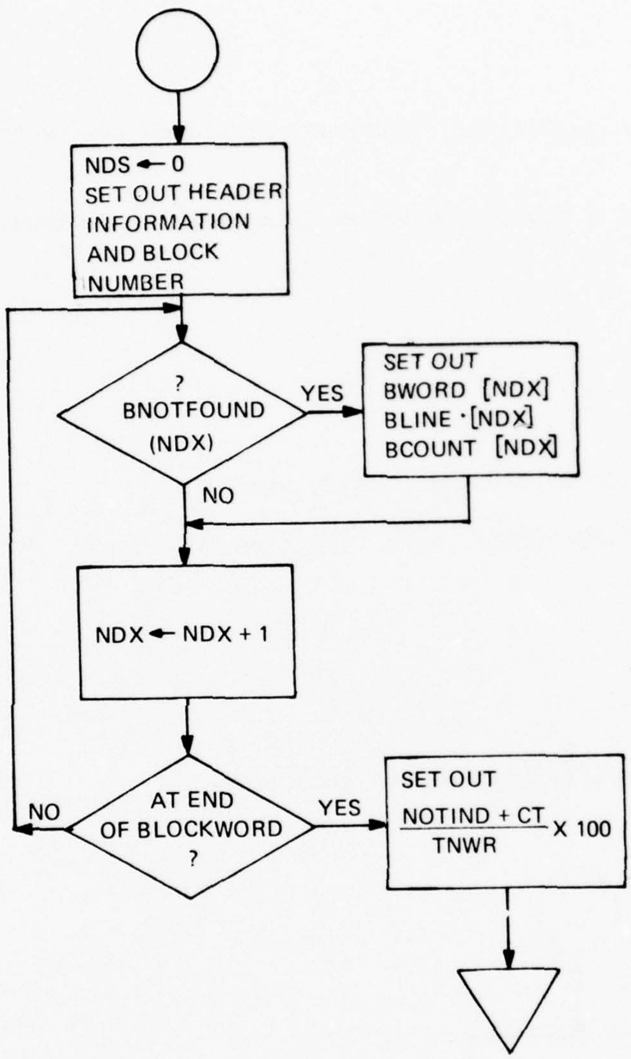
CALLING MODULES: None

COMMENTS: The norm data must be included in this module. The BLOCKSUM is a simple computation module. The exact interpolation method can be determined later.


NSNB ← 0
TNWR ←
TNWR + TNWB
LB ← (1 − NCLB/TNSB)
RB ← (TNSB/(TNSB + NCRB))
SE ← (1 − NNPB/TNSB)
FOR EACH ENTRY IN BLOCKWORD TABLE
? BMULTY TRUE
YES
NSNB ← NSNB + 1
NO
END OF TABLE
FOR EACH MEASURE M
INTERPOLATE VALUE OF M USING NORMTABLE [M,TY] WHERE TY IS TYPE OF TEXT (SEE NORM. . .BOOLEANS)
UPDATE TOTALS OF ALL MEASURES;
CALCULATE COMPOSITE INDEXES BASED ON VALUES OF READER HIGH, READER HI LOW, READER LOW AS DESCRIBED UNDER TECHNIQUE

PROGRAM MODULE SPECIFICATION

NAME:	CHECKOUT
NUMBER:	17
PURPOSE:	CHECKOUT displays the block results of the dictionary check.
TECHNIQUE:	CHECKOUT is a report generator using values previously calculated.
INPUT:	None
FILES ACCESSED:	None
GLOBAL DATA:	Uses BLOCKWORD TABLE and TNWR and NOTINDICT.
OUTPUT:	The output of CHECKOUT is a report on the results of the dictionary check.
MODULES CALLED:	None
CALLING MODULES:	None
COMMENTS:	None


NDS ← 0
SET OUT HEADER INFORMATION AND BLOCK NUMBER
? BNOTFOUND (NDX)
YES
SET OUT BWORD [NDX] BLINE ·[NDX] BCOUNT [NDX]
NO
NDX ← NDX + 1
AT END OF BLOCKWORD ?
NO
YES
SET OUT NOTIND + CT / TNWR X 100

NAME: MEASUREOUT

NUMBER: 18

PURPOSE: MEASUREOUT displays the measures calculated for each block.

TECHNIQUE: MEASUREOUT is a report generator using values previously calculated.

INPUT: None

FILES ACCESSED: OUTPUT FILES

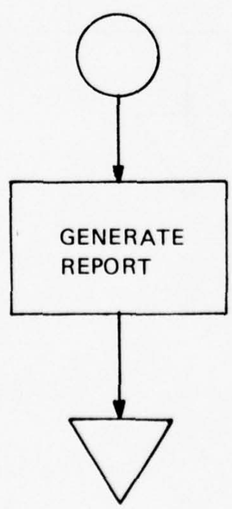
GLOBAL DATA: Uses all measures previously calculated.

OUTPUT: The output of MEASUREOUT is a report on the block's measures.

MODULES CALLED: None

CALLING MODULES: None

COMMENTS: MEASUREOUT is a trivial report generator.



NAME: RUNSUM

NUMBER: 19

PURPOSE: The RUNSUM module summarizes the results of the run in preparation for the RUNOUT report module.

The RUNSUM module updates the dictionary header information to reflect the current run.

TECHNIQUE: RUNSUM takes totals calculated by BLOCKSUM at the end of each block and computes averages.

INPUT: Totals from BLOCKSUM.

FILES ACCESSED: None

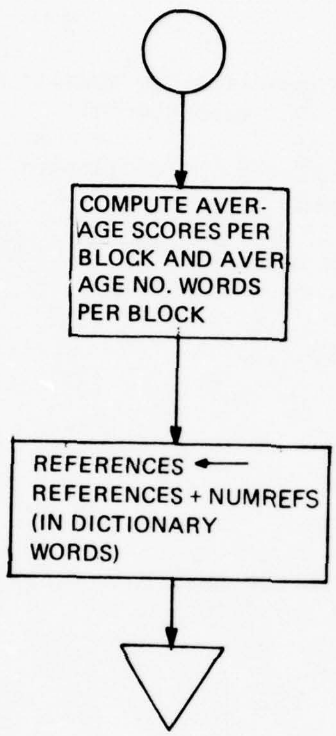
GLOBAL DATA: Totals from BLOCKSUM.

OUTPUT: The output of the RUNSUM module is a set of averages to be used in the RUNOUT module.

MODULES CALLED: None

CALLING MODULES: None

COMMENTS: The RUNSUM module is a simple calculation module.


COMPUTE AVERAGE SCORES PER BLOCK AND AVERAGE NO. WORDS PER BLOCK
REFERENCES ← REFERENCES + NUMREFS (IN DICTIONARY WORDS)

NAME: RUNOUT

NUMBER: 20

PURPOSE: RUNOUT displays the final run results.

TECHNIQUE: RUNOUT is a report generator using values previously calculated.

INPUT: None

FILES ACCESSED: OUTPUT FILE

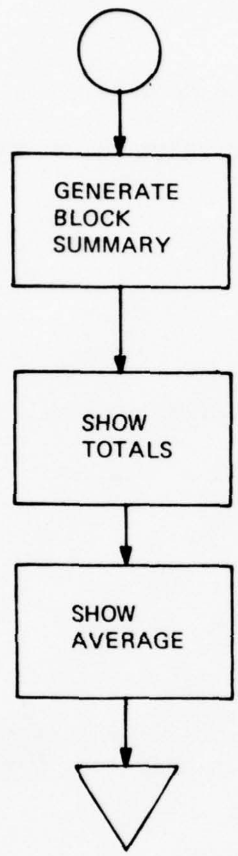
GLOBAL DATA: Uses data calculated at end of each block.

OUTPUT: The output of RUNOUT is a final run report. See Figure D-4 for format and content.

MODULES CALLED: None

CALLING MODULES: None

COMMENTS: RUNOUT is a report generator using values calculated at the conclusion of each block. In practice, RUNOUT may be called for each block.


GENERATE
BLOCK
SUMMARY
SHOW
TOTALS
SHOW
AVERAGE

APPENDIX C

Run Request Syntax

RUN REQUEST SYNTAX

Appendix C defines the format, content, and sequence for a user of the CM program to enter run requests into the computer. This information, termed run request syntax is specified in the form of a series of syntax diagrams which will apply whether the input is prepared on punched cards and read via a local or remote card reader or is entered via a keyboard on a local or remote interactive computer terminal.

The syntax diagram was selected to form the basis for this appendix because it affords a concise exposition of a syntax involving defaults, alternatives, and iterations; it is rigorous without being cumbersome. There are few formal rules: The basic rule is that any path traced along the forward direction of the arrows will produce a syntactically correct component of the run request language.

There are two kinds of diagram components: terminals & non-terminals Non-terminals are items which have their own tracks and are indicated by < >. Terminals are items which do not have their own tracks. These are either special symbols or "words." Words which are entirely underlined must be completely specified. Words partially underlined can be abbreviated by the underlined part or any part of the entire word containing at least the underlined part. Thus, → SENTENCES → can be represented by any of

SENT   SENTENCE   SENTEN

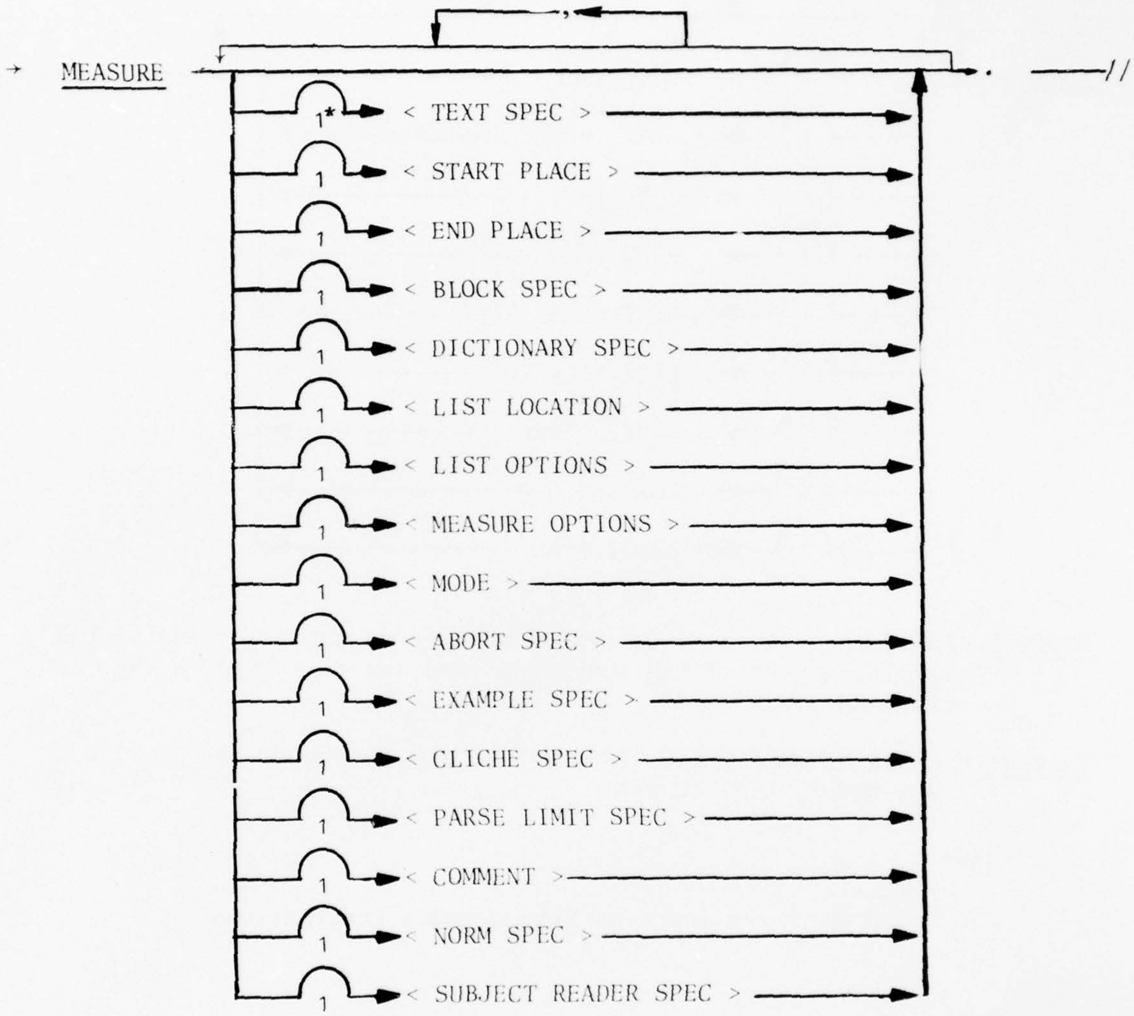
but   SEN   SENTNC   SENTENCESX

are all unacceptable.

Iteration is noted by —n→ in a track the horizontal directed line containing the —n) can be passed at most n times (but can be passed 0 times). If an asterisk (*) follows the number, then the containing directed horizontal line must be passed at least once.

Basically, two types of requests are permissable, those requesting a check of text against the dictionary, (COMMAND: CHECK) and those requesting a calculation of comprehensibility measures (COMMAND: MEASURE).

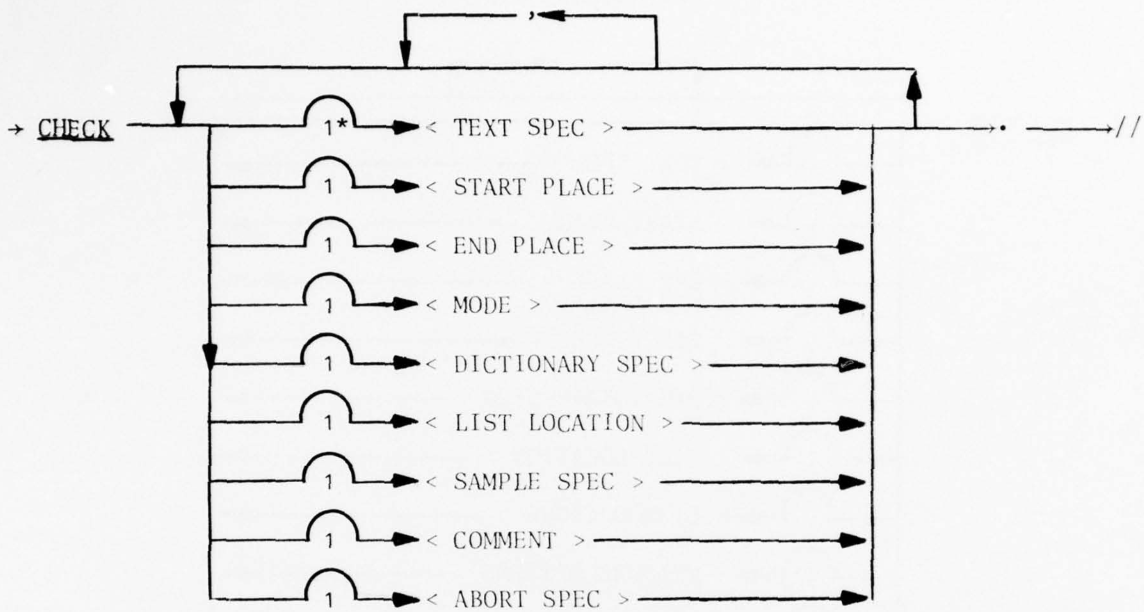
RUN REQUEST SYNTAX



PURPOSE • Initiates calculation of comprehensibility measures for a specified part of a specified file. Each syntax element is elaborated on in subsequent sections.

EXAMPLE • MEASURE USING FILETWO FROM LINE 1000 THRU END, BLOCK SIZE 50 SENT, LIST TO PRINTER TEXT, COMP SUMMARY, INTERACTIVE.

REQUEST RUN SYNTAX



PURPOSE • Initiate a check to determine whether or not all words in the specified part of the specified file are contained in the dictionary.

Example • CHECK USING FILEONE, FROM START, FOR 14 BLOCKS, BLOCK SIZE 400 WORDS, LIST TO REM.

< FILE ID > ::= {name of file & media information}

PURPOSE • Defines the name and any necessary media information of an input file (text, dictionary, etc.)

COMMENTS • The exact form of a < FILE ID > will depend on the particular implementation system and language.

< TEXT SPEC > ::=

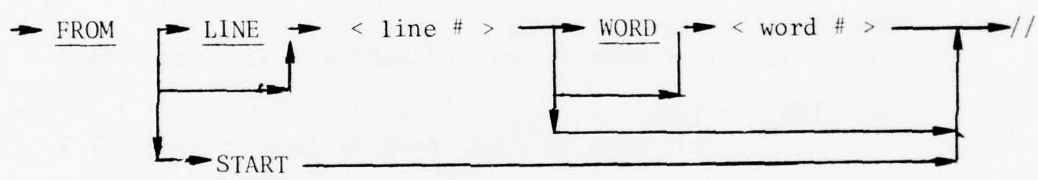
→ USING → < FILE ID > → //

PURPOSE • Defines the input text to be processed.

COMMENTS • The text file will consist of fixed-length blocks containing fixed-length records. The file will consist of two parts: header and text. The header portion will consist of information about the text file including text title and author, classification, creation and revision dates and versions. The text portion will contain the actual text, one line per record.

The exact format of the header records and the file's record and block sizes will be defined during the program development.

< START PLACE > ::=



PURPOSE • Defines starting location in file.

COMMENTS • FROM START: start at beginning of file (line 1 word 1)

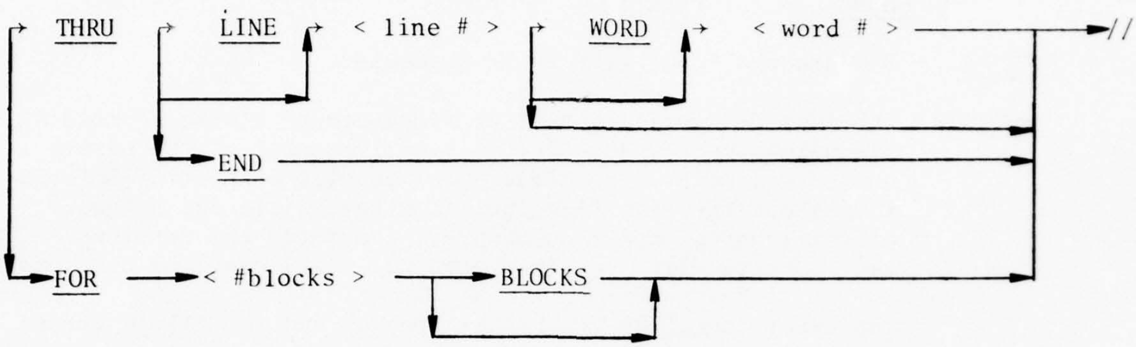
FROM LINE <line # > WORD < WORD # > :

start at word < word # > on line < line # >
<line # > $\geq$ 1, < word # > $\geq$ 1.

FROM LINE < line # > :
start at word 1 on line < line # >

If no < start place > is provided, < FROM START > is assumed.

< END PLACE > ::=



PURPOSE • Define ending location in file.

COMMENTS • THRU END: to end of file

THRU LINE: < line # > WORD < word # > :
last word is word < word # > on line < line # > .

THRU LINE: < line # > :
last word is final word on line < line # > .

FOR < # blocks > BLOCKS : < # blocks > text blocks
will be processed.

If no < end place > is provided, THRU END is assumed.

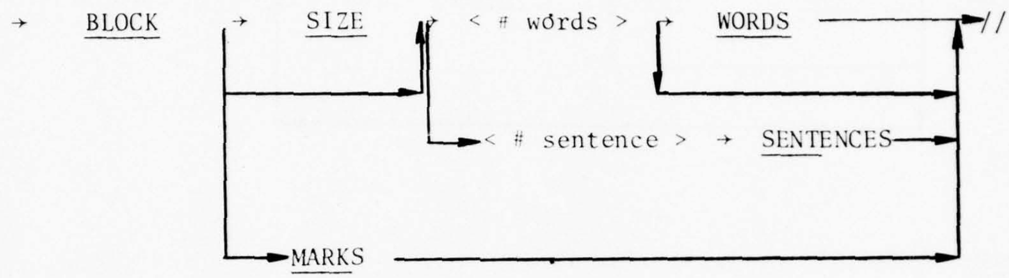
< DICTIONARY SPEC > ::=

→ DICTIONARY → < FILE ID > →//

PURPOSE • Specifies the dictionary to be used.

COMMENTS • If no < dictionary spec > is provided, the standard dictionary will be used. The format of dictionary files and the name of the default dictionary will be specified during the development phase.

< BLOCK SPEC > ::=



PURPOSE • Specifies the size of a text block.

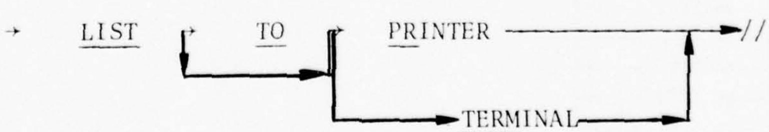
COMMENTS • BLOCK SIZE < # words > WORDS : blocks will be formed at least < # words > words such that a block contains only complete sentences. 100 ≤ < # words >

BLOCK SIZE < # sentences > SENTENCES : blocks will be formed of < # sentences > sentences.

BLOCK MARKS : the source will define the blocks by means of embedded < block mark > (depending on particular system & text editor it may be necessary to specify the < block mark > in the request syntax).

If no < block spec > is provided, BLOCK SIZE 500 WORDS is assumed.

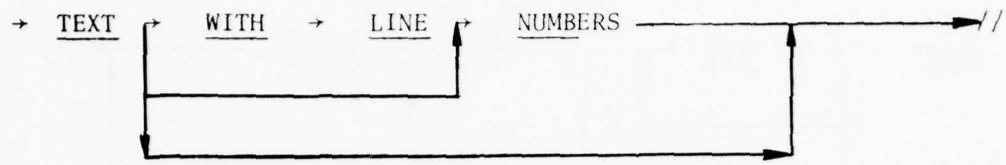
< LIST LOCATION > ::=



PURPOSE • Specifies location of output display or listings.

COMMENTS • TERMINAL is the originating remote station/terminal. If no < list location > is specified, LIST TO PRINTER is assumed.

< LIST OPTION > ::=

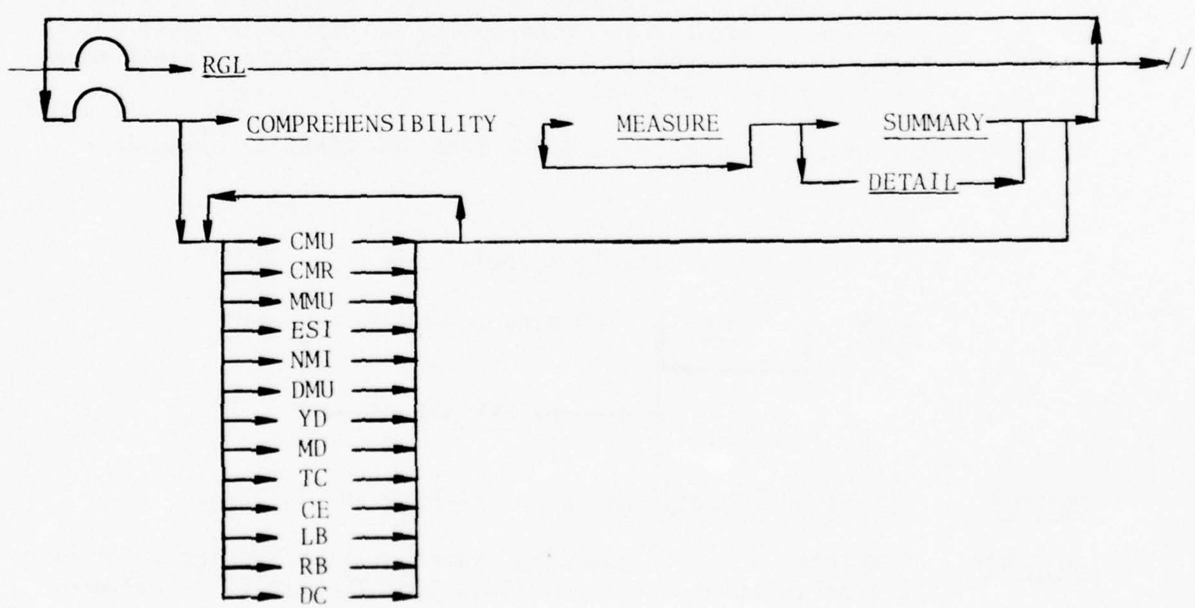


PURPOSE • Specific optional list items.

COMMENTS • If specified, the text blocks reported on will be listed. If "WITH LINE NUMBERS" is included, the line numbers will be printed with each text line.

If no < list option > is specified, no text lines will be printed.

< MEASURE OPTION > ::=



PURPOSE • Specifies what measures are to be calculated.

COMMENTS • If no <measure option> is specified, COMP SUMMARY is assumed.

CMU - Cognition of Semantic Units
CMR - Cognition of Semantic Relations
MMU - Memory of Semantic Units
ESI - Evaluation of Symbolic Implications
NMI - Convergent Production of Semantic Implications
DMU - Divergent Production of Semantic Units
YD - Yngve depth
MD - Morpheme depth
TC - Transformational complexity
CE - Center embedding
LB - Left branching
RB - Right branching
DC - Deleting complements

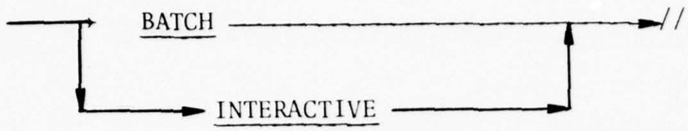
There are five types of report that may be provided:

1. Sentence detail
2. Dictionary check
3. Block results
4. Run summary
5. RGL report

Only report 2 is provided on a "CHECK" run. For a "MEASURE" run:

Report 1 is provided if "DETAIL" is specified.
Report 3 is provided if "SUMMARY" is specified.
Report 4 is always provided.
Report 5 is provided if "RGL" is specified.

< MODE > ::=

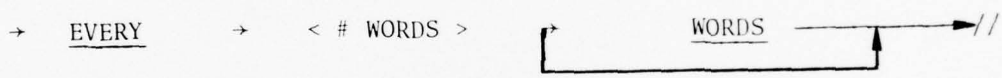


PURPOSE • Activates batch processing.

COMMENTS • If BATCH,the run can not respond to any system questions or notifications. If interactive,the user can respond to system questions or notifications (such as word not in dictionary).

If no < mode > is specified, BATCH is assumed.

< SAMPLE SPEC > ::=



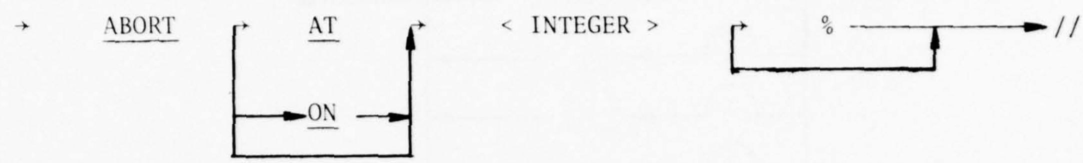
PURPOSE • Specifies that the dictionary check is to be performed only for one out of each < # words > words.

COMMENTS • < # words > an integer greater than 1.

The percentage of words not found in the dictionary will be reported.

If no < SAMPLE SPEC > is provided, each word will be checked, i.e., < # words > will be set to 1.

< ABORT SPEC > ::= .



PURPOSE • The abort analysis if current dictionary is incomplete.

COMMENTS • If specified, the entire run will be aborted if more than < integer > percent of the words are not in the dictionary. < integer > between 0 and 100 inclusive. The abort action will be considered only at the end of each text block. If the percentage of words not in the dictionary at the end of a text block exceeds < integer > , all measures and statistics will be calculated and reported thru the aborting block, including all summary information for the run. No additional text blocks will be processed.

If no < ABORT SPEC > is provided, all text blocks will be processed regardless of the number of "unidentified" words.

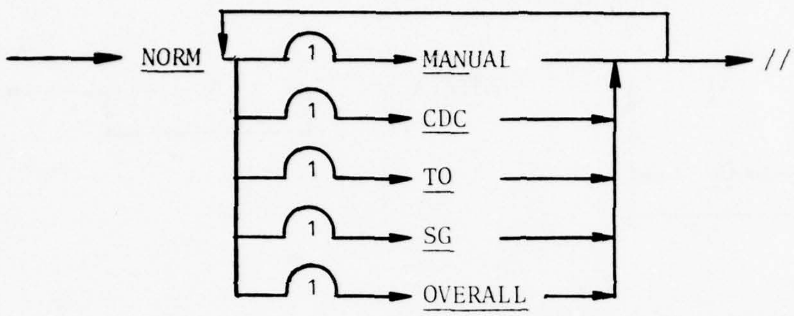
< EXAMPLE SPEC > ::=

PURPOSE • Specifies the name of the example file.

COMMENTS • This file contains all "for example" words and phrases required to compute the Divergent Production of Semantic Units (DMU) measure. If no < example spec > is provided, the standard example file will be used.

The format of the example file and the name of the default example file will be specified during the development phase.

< NORM SPEC > ::=

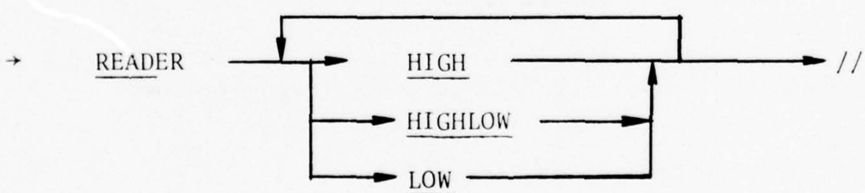


PURPOSE • Specifies the norm(s) to which the current run is to be compared.

COMMENTS • MANUAL - manual
CDC - Career Development Course
TO - Technical Order
SG - Study Guide
OVERALL - Overall

If no < norm spec > is provided, OVERALL is assumed.

< SUBJECT READER SPEC > ::=



PURPOSE • Specifies the subjects reading skill level(s).

COMMENTS • One or more levels may be selected. The subject reader specification value(s) will determine which composite index formula(s) will be used.

If no < subject reader spec > is provided, all three are assumed.

< CLICHE SPEC > ::=

→ CLICHE → < file id > ———→ //

PURPOSE • Specifies the name of the cliche file.

COMMENTS • If no < cliche spec > is provided, the standard cliche file will be used.

The format of the cliche file and the name of the default cliche file will be specified during the development phase.

< PARSE LIMIT SPEC > ::=

→ PARSE → MAX → < integer > ———→ //

PURPOSE • Specifies the maximum number of parses that will be permitted per sentence. < integer > between 1 and 100 inclusive.

COMMENTS • If no < PARSE LIMIT SPEC > is provided, no limit will be placed on the number of possible parses per sentence.

< COMMENT >

COMMENT → " → < any string not containing a " > →" ——→//

PURPOSE • Specifies a comment or title that will appear at the top of each printed report page for this run. The string can be any series of printable characters not containing a quote ("). The entire string and its surrounding quotes must be on the same input record.

EXAMPLE:

COMMENT "PRELIMINARY REFERENCE MANUAL #2"

COMMENT • If no comment is provided the following title is assumed:

"Comprehensibility Measures Program"

APPENDIX D

Output Formats

< REPORT TITLE FROM COMMENT REQUEST >

DETAIL RESULTS FOR XXX SENTENCES, BLOCKS XXX TO XXX OF FILE < text file id >, LINE NUMBERS XXXXX TO XXXXX
RUN DATE: XX-XX-XX RUN TIME: XX: XX: XX FILE DATE: XX-XX-XX

Sentence No.	Block No.	First Word		Number per Sentence										
		Line	Word	Words	Potential Parses	Explan-ation	Yngve Depth	Trans Complex	Noun Phrases	Modify Clauses Right	Modify Clauses Left	Deleted Complem.	Mor-phemes	Parts Speech
				TNWS	NPPS	NESS	YDS	TCS	NNPS	NCRS	NCLS	DCS	NO MORE	NO PARTS
001	001	XXX	XX	XXXXX	XX	XX	X.XX	X.XX	XX	XX	XX	XX	XXX	XXX
002	001													
Total	001	-	-											
001	002													
.														
.														
Total	-													
Gr. Total	-	-	-	XXXXXX	XXX	XXX	XX.XX	XX.XX	XXX	XXX	XXx	XXX	XXXX	XXXX
Average per sentence over all blocks		-	-	XXX	XX	XX.X	X.XX	X.XX	XX.X	X.X	X.X	X.X	XX.X	XXX.X

Figure D-1. Format of output for sentences.

Source: Program Module SENTOUT (module 13)

< REPORT TITLE FROM COMMENT REQUEST >

DICTIONARY CHECK RESULTS FOR FILE < text file id > LINE NUMBER XXXXX TO XXXXX
RUN DATE: XX-XX-XX, RUN TIME XX:XX:XX FILE DATE: XX-XX-XX

START AT WORD XX ON LINE XXX THRU { WORD XX ON LINE XXX / XXX BLOCKS / END OF TEXT } SAMPLE EVERY XX WORDS

Block No.	Word not in Dictionary	First Occurrence (Line No.)*	Total No. of Occurrences in block	% Words not in Dictionary	Analyst/Editor Comment 1. spelling change 2. add to dictionary 3. dictionary change: Code	Word	Dictionary Contents
001	XXXXXXXXXXXXXXX	XXX	XXXX	-			
	. . .						
001			XXXX	XX.XX			
002							
. . .							
RUN TOTAL	-	-	XXXX	XXXX.XX			
AVERAGE				XX.XX			

*List of words in each block is printed in alphabetical order and line no. of 1st. occurrence.

Figure D-2. Format of output for check of words in dictionary.

Source: Program Module CHECKOUT (module 17)

< REPORT TITLE FROM COMMENT REQUEST >

RESULTS FOR BLOCK NO. XXX OF FILE < text file id > LINE NUMBERS XXXXX TO XXXXX
RUN DATE: XX-XX-XX, RUN TIME: XX:XX:XX, FILE DATE: XX-XX-XX
FROM LINE XX, BLOCK XXXX TO LINE XX, BLOCK XX FOR XXXXX WORDS

Structure-of-Intellect			Psycholinguistic			Other Block Results			
Measure	Value	Normalized	Measure	Value	Normalized	Variable	Value	Variable	Value
CMU			YD			NDWB		NPPB	
CMR			MD			TNWB		TNEB	
MMU			TC			NAMB		TNSB	
ESI			SE			NASB		OSWB	
NMI			LB			NSNB		TNCB	
DMU			RB			NORB		TSCB	
			DC			NDNB		TLCB	
						NSWB		TPSB	
						TNMB			
						NWNDB			

READER SKILL LEVEL	:	HI	HI & LO	LO
COMPOSITE VALUE OF MEASURES	:	XXX.XX	XXX.XX	XXX.XX

Figure D-3. Format of output for measures in a single block.

Source: Program Module MEASUREOUT (module 18)

< REPORT TITLE FROM COMMENT REQUEST >

RUN SUMMARY FOR FILE ID, BLOCK SIZE IS XXXXX {WORDS / SENTENCES} FROM WORD XX LINE XXXX THRU {WORD XX LINE XXXXX / END OF TEXT}

INPUT REQUEST: {AS GIVEN BY USER}
RUN DATE: XX-XX-XX, RUN TIME: XX:XX:XX, FILE DATE: XX-XX-XX

DICTIONARY < FILE ID >, PARSE LIMIT XX NORM < norm list > SUBJECT READER < skill levels >

M-Measured N-Normalized		AVERAGE PERCENTILE MEASURES OVER THE BLOCK												Number of					First Word	
	Block	Structure-of-Intellect					Psycholinguistic				Composite			Words	Sentences		Parses/	Morphemes		
	No.	CMU	CMR	MMU	...	Etc.	YD	MD	...	Etc.	HI Skill	HI&LO Skill	LOW Skill	TNW	Parse	NP	Sentence	per Word	Line	Word
M	XXXX	X.XX	X.XX	X.XX		X.XX	X.XX	X.XX		X.XX	–	–	–	XXXXX	XXX	XX	XX.X	XXX.XX	XXXX	XX
N	XXXX	XX.X	XX.X	XX.X	...	XX.X	XX.X	XX.X		XX.X	XXX.X	XXX.X	XXX.X	–	–	–	–	–	–	–
M																				
N																				
M	Repeat for each block																			
N																				
M	Avg/ Block	XXX.XX	XXX.XX	XXX.XX		XXX.XX	XXX.XX	XXX.XX		XXX.XX				XXXXX	XXX	XX	XX.XX	XX.X	–	–
N	Avg/ Block	XX.X	XX.X	XX.X		XX.X	XX.X	XX.X		XX.X				–	–	–	–	–	–	–
M	Sigma/ Block													–	–	–	XX.X			

Figure D-4. Format of output listing results for an entire run.

Source: Program Module RUNOUT (month 20)

< REPORT TITLE FROM COMMENT REQUEST >

READING GRADE LEVEL RESULTS FOR FILE < text file id > , BLOCKSIZE: XXXX WORDS
RUN DATE XX-XX-XX, RUN TIME: XX:XX:XX, FILE DATE: XX-XX-XX

Block No.	READING GRADE LEVEL VALUES			Automated Readability Index	FIRST WORD		NUMBER OF	
	FORECAST	AUTOMATED READABILITY INDEX	FLESCH		LINE	WORD	LINE	WORDS
XXXX	XX.XX	XX.XX	XX.XX	XX.XX	XXXX	XX	XXX	XXXXX
Total	XXX.XX	XXX.XX	XXX.XX	XXX.XX	-	-	XXXXX	XXXXX
Avg./ Block	XX.XX	XX.XX	XX.XX	XX.XX	-	-	XXX	XXXXX

Figure D-5. Format for output listing results for RGLs for a run.

Source: Program Module RGLOUT (module 15)

APPENDIX E

Names of Global and Other Data Items

Each item in Appendix E is shown in the form:

<name> <type> <initial value> <reference list> <description>

<name> is the name of the item as used in the program module descriptions. In actual implementation, name changes may be required to correspond to language requirements.

<type> is one of

- F file
- B boolean, either true or false
- D discreet, can have one of a list of discreet values. Followed by the possible values. Note, a discreet item is never true or false, and can never be used in an arithmetic calculation. It can only be set to one of its possible values and compared to one of its possible values.
- R real
- A alpha (string of characters)
- I pointer or index

If followed by ARRAY then item is an array of items.

<initial value> is indicated by a value in square brackets, i.e., [0]. this is the value the item is set to at run initialization. If no initial value is specified, its initial value can be anything. Initial values are set in module 1. An initial value of [∞] implies the largest possible integer value.

<reference list> Each module that references an item has its module number in this list. If followed by an asterisk "*" then the value of the item may be changed in that module. The module "C" denotes the common control logic outside of any specific module.

<description> is the purpose of the item.

Files

CLICHEFILE F 2,7
list of cliches

DICTFILE F 2,7
dictionary

EXAMPLEFILE F 2,7
list of example introducer words and phrases

INPUT FILE F 1,4
user input file (card or remote).

REPORTBLOCK F 2,7,18
printer file for block summary.

REPORTCHECK F 2,17
printer file for check run summary.

REPORTRGL F 2,15
printer file for RGL summary.

REPORTSENTENCE F 2,13
printer file for sentence summary.

REPORTSUMMARY F 2,5,20
printer file for run summary.

TEXTFILE F 2,3,7
text file.

General Variables

ABORTPERCENT R [100] 2*, C
percentage of words not in dictionary to total words that is to cause a run to abort. See <abort spec>.

ARI R 14*,15,18
Automated Readability Index

ARIRGL R 14*,15,18

Automated Readability Index Reading Grade Level

BLOCKCOUNT R 2*, 13, C

number of blocks to process. See < end place > .

BLOCKNUMBER R [C] 3*
number of block being processed.

BLOCKSIZE R [500] 2*,17,20,C
size of each block (units determined by BLOCKTYPE). See <block spec>.

BLOCKTYPE D: BLOCKINWORDS/BLOCKINSENTENCES/BLOCKMARKS
[BLOCKINWORDS] 2*,17,20,C
indicates the units of block length. See <block spec>.

COMMENT A 2*,13,15,17,18,20
holds comment string for page header. See <comment request>.

COMPINDEXHIGH
COMPINDEXHILOW R 16*, 20
COMPINDEXLOW
contains composite index based on reader skill.

CURRENTLINE R 3*
current text line number in progress.

ENDLINE R [∞] 2*,17,20,C
last line to be processed. See <end place>.

ENDTYPE D THRUEND/COUNTEND [THRUEND] 2*,20,C
type of <end place>. If THRUEND then ENDLINE and ENDWORD are valid. If COUNTEND then BLOCKCOUNT is valid. See <end place>.

ENDWORD R [∞] 2*, 17,20,C
last word to process on ENDLINE. See <end place>.

ERRORCOUNT R [0] 5*
number of syntax errors in a run.

FLESCHRL R 14*,15,18
Flesch Reading Grade Level.

FORCASTRGL R 14*,15,18
Forcast Reading Grade Level.

FOUND B 7*,8
true if current word is in dictionary.

HAVEERROR B [false] 2,5*
true when have a syntax error on a request.

LISTLINE B [false] 2*,7
true if line numbers are to be included in text list.
See <list options>.

LISTLOC D LISTTOPRINTER/LISTTOREMOTE [LISTTOPRINTER]

.2*,7

where to list output reports. See <list location>.

LISTTEXT B [false] 2*,7
true if to list text in block summary report. See <list options>.

MAXPARSE R [∞] 2*,20
maximum number of parses per sentence permitted. See <parse limit spec>.

MEASURE *B array [true] 2*,10,11*
one item for each measure: true if to compute that measure. See <measure options>.

MODE D BATCH/INTERACTIVE 2*,7
indicates whether the user is running the program from batch or interactively. See <mode>.

NDNSAVE R 3*,8*,10
save value of NDNB when block goes over 100 words.

NDWSAVE R 3*,8*,10
save value of NDWB when block goes over 100 words.

NORMCDC	(Career Development Course)	B	[false]	2*,20
NORMMAN	(manual)		[false]	
NORMOVERALL	(overall)		[true]	
NORMSG	(study guide)		[false]	
NORMTO	(Technical Order)		[false]	

true if norm is to be calculated on specified basis. See <norm spec>.

NOTINDICT R [0] 7*,17
total number of words not in dictionary.

NUMREFS R [0] 7*
number of references made to dictionary in this run.

PARSE-TABLE array 9*,11
holds parse information for current sentence.

READERHIGH
READERHILOW B [true] 2*,16,20
READERLOW
true if the reading skill of the subject is to be high, high/low, or low. See <subject reader spec>.

RUNTYPE D CHECKRUN/MEASURERUN 2*,C
indicate type of run.

SAMPLESIZE R [1] 2*,17
for CHECK run only: perform dictionary check on one of every SAMPLESIZE words. See <sample spec>.

SCANINTOKEN A 2,4*,5,7
current token from SCANINPUT.

SANINTYPE D TOKENISWORD/TOKENISNUMBER/TOKENISFILED/TOKENISSYMBOL/ TOKENISEOF 2,4*,7
type of token returned from SCANINPUT.

SCANINVALUE R 2,4*,7
if token returned from SCANINPUT is a number (SCANINTYPE= TOKENISNUMBER), then SCANINVALUE is the integral value of the current token.

SCANPOINTER I [81] 4*
pointer to current column in SCANRECORD.

SCANRECORD A array 4*
current input record

SCANSAMETOKEN B [false] 2*,4*,7*
if true, SCANINPUT will look at the same token as the last call. Always FALSE after a call on SCANINPUT.

STARTLINE R [1] 2*,3,17,20
starting text line. See <start place>.

STARTWORD R [1] 2*,3,17,20
starting text word on STARTLINE. See <start place>.

Run Level

TNWR R [0] 8*,17
total number of words processed

Block Level

NDWB R 3*,8*,10,18
number of different words in a text block

TNWB R 3*,8*, 10,14,18
total number of words in a text block

NAMB R 3*,8*,18
number of multiple word abbreviations in a text block

NASB R 3*,8*,18
number of single word abbreviations in a text block.

NSNB R 3*,8*,10,16*, 18
number of shared nouns in a text block

NORB R 3*,8*,10,18
number of references in a text block

NDNB R 3*,8*,10,18
number of different nouns in a text block

NSWB R 3*,8*,10,18
number of abbreviated or symbolic words in a text block

TNSB R 3*,6*,10,11*,13,14,18
number of sentences in a text block

TNMB R 3*,8*,18
number of morphemes in a text block

NPPB R 3*,9*,10,18
number of potential parses in a text block

TNEB R 3*,8*,10,18
number of elucidations in a text block

NWNDB R 3*,7*,18
number of words not in dictionary in a text block

TPSB R 3*,8*,18
number of parts of speech of all words in a text block

TCLB R 3*,8*,18
number of clauses on left of noun in a text block

TNCB R 3*,8*,14,18

number of characters in a text block

OSWB R 3*,8*,14,18

number of one syllable words in a text block

TSCB R 3*,8*,14,18

number of syllables in all words in a text block

NCRB } R [0] 12*,16
NCLB }

number of modifying clauses on the right/left of the object noun for a block

Sentence Level

NPPS R 6*,13

number of potential parses in a sentence

NESS R 6*,13

number of explanations in a sentence

TNWS R 6*,7*,11,13

number of words in a sentence

YDS R 6*,11*,13

Yngve depth of a sentence

TCS R 6*,13

Transformational Complexity of a sentence

NNPS R 6*,10*,12,13,16

number of noun phrases to the right of the subject verb in a sentence

NCRS } R 6*,13
NCLS }

number of modifying clauses on the right/left of the object noun phrase of a sentence

DCS R 6*,13

Deleted Complement of a sentence

Measures

For each measure there are three items.

1. The measure itself (a two or three-character name), e.g.: CMU.
2. A normalized value of the measure (measure followed by "N"), e.g., CMUN.
3. A total measure used to compute averages (measure followed by "T"), e.g., CMUT.

All items are of type R.

The measures are divided into two categories:

<u>Structure-of-intellect</u>: (referenced in 10*,16,18,20)

CMU - Cognition of Semantic Units
CMR - Cognition of Semantic Relations
MMU - Memory of Semantic Units
ESI - Evaluation of Symbolic Implications
NMI - Convergent Product of Semantic Implications
DMU - Divergent Product of Semantic Units

<u>Psycholinguistic</u>: (referenced in 11*,16,18,20)

YD - Yngve Depth
MD - Morpheme Depth
TC - Transformational Complexity
SE - Center Embedding
LB - Left Branching
RB - Right Branching
DC - Deleted Complement

The 13 normalized values (CMUN, CMRN, MMUN, ESIN, NMIN, DMUN, YDN, MDN, TCN, SEN, LBN, RBN, DCN) are referenced in 16*, 18, 20.

The 13 total values (CMUT, CMRT, MMUT, ESIT, NMIT, DMUT, YDT, TCT, SET, LBT, RNT, DCT) are referenced in 18*, 20.

Sentence Data Array (one element of each for each word in a sentence)

NOMORE	R	7*,8,9	number of morphemes
NOPARTS	R	7*,8,9	number of parts of speech
PART[*]	R	array 7*,9	parts of speech (one item for each part of speech)
NOSYLLABLES	R	7*,8,9	number of syllables
NEGIND	B	7*,9	negative indicator
SYMBOLIND	B	7*,8,9	symbolic/abbreviation indicator
NOWORDS	R	7*,8,9	number of words
NOREFS	R	7*,8	number of references to word
WORD	A	7*	word
EXAMPLE	B	7*	example indicator
CLICHE	B	7*	cliche indicator

Dictionary File Items (items have same meaning as corresponding item in SENTENCE DATA ARRAY)

DNOMORE	R	7
DNOPARTS	R	7
DPART [*]	R	array 7
DNOSYLLABLES	R	7
DNEGIND	B	7
DSYMBOLIND	B	7
DNOWORDS	R	7
DNOREFS	R	7*
DWORD	A	7

header information:

DICTDATE	R		dictionary creation date
TOTALPARTS	R		total number of parts of speech
REFERENCES	R	19*	total number of references to dictionary

BLOCKWORD TABLE (one entry for each different word in a block)

BWORD	A	8*, 17	word
BNON	B	8*,17	noun indicator
BNOTFOUND	B	8*,17	time if word not in dictionary
BLINE	R	8*,17	line number of first occurrence if BNOTFOUND is true
BCOUNT	R	8*,17	number of times word used if not in dictionary
BMULTY	B	8*,17	true if used more than once in a text block

Appendix F

Partial Set of Phrase Structure Parsing Rules
Sentence Element Category Symbols

F DEFINITIONS/ABBREVIATIONS FOR SENTENCE ELEMENT CATEGORY (PARTS OF SPEECH AND PHRASE STRUCTURE) SYMBOLS

S	sentence
VP	verb phrase
NP	noun phrase
M	modifier
N	noun
PRTP	participial phrase
PREPP	prepositional phrase
PP	past participle
V	verb
LV	linking verb (is)
P	preposition
CONJ	conjuntion
NEG	negative
PRN	pronoun
RP	relative pronoun
MG	modifier group
ADV	adverb
DET	determiner
PRT	participle
EXP	introducer of explanation
NABBR	noun abbreviation
AABBR	adjective abbreviation
PN	proper noun
APOSP	apposite phrase
NPC	noun phrase complement
VPC	verb phrase complement
COM	complementizer
PART	particle
INFP	infinitive phrase
INF	infinitive
AUX	auxiliary
RECL	relative clause
ADVCL	adverbial clause
ADVP	adverbial phrase
CL	clause
C	conjoint
SA	sentence taking adverb
MP	modifier phrase
PURPCL	purpose clause
VB	verbal

F SOME PHRASE STRUCTURE PARSING RULES

	ADVCL	
ADVCL	→	CONJ S
ADVCL	→	ADV S
ADVCL	→	NP ADV S
	ADVP	
ADVP	→	CONJ PARTP
	AUX	
AUX	→	AUX NEG
	C	
C	→	CONJ S
C	→	"to" S (the word "to")
	CONJ	
CONJ	→	M ADV
	INFP and INF	
INFP	→	INFP VP
INF	→	"to" LV
	N	
N	→	PRN
	NP	
NP	→	NP PP
NP	→	N MP
NP	→	DET NP
NP	→	DET N
NP	→	M N
NP	→	NP CONJ
NP	→	CONJ NP
NP	→	N CONJ N
NP	→	N NP
NP	→	M NP
NP	→	N CONJ N
NP	→	N RELCC
NP	→	MP N
NP	→	CONJ NP
	NPC	
NPC	→	COM S
	PRTP	
PRTP	→	PRTP PP
PRTP	→	PRT NP

PP

```
PP        →    PP CONJ PP
PP        →    P NP
PP        →    PP PP
PP        →    P N
```

S

```
S         →    ADVCL S
S         →    NP VP
S         →    PURPOSECL S
S         →    S PURPOSECL
S         →    S CONJ S
        or S → S CONJOINT
        CONJOINT → CONJ S
S         →    PP S
S         →    CONJP S
S         →    S NPC
S         →    VP VPC
S         →    CL S
S         →    PP VP
S         →    C S
S         →    PRTP S
S         →    VP PP
S         →    VP
S         →    S CONJ S
S         →    VP CL
```

VP

```
VP        →    V NP
VP        →    VB NP
VP        →    V NP PNP
        or VP →  VP PREPP
           NP →
        or NP →  NP PREPP
           PREPP →  P NP
           VP →  V NP
VP        →    VP VPC
VP        →    ADV PRT
VP        →    ADV VP
VP        →    VP ADV
VP        →    VB
VP        →    LV M
VP        →    VB ADV
VP        →    AUX VP
VP        →    VP VB
VP        →    VB NEG
VP        →    LV VB
VP        →    LV PRT
VP        →    VP PP
VP        →    VP VPC
VP        →    VP CONJ VP
VP        →    ADVP VP
VP        →    VB NP
```

APPENDIX G

Norms for Each Measure

1. Cognition of Semantic Units (CMU)

Percentile	SG	MAN	CDC	TO	O/A *
90	.470	.436	.462	.534	.476
80	.450	.407	.442	.496	.449
70	.442	.390	.403	.458	.423
60	.424	.366	.374	.434	.400
50	.394	.340	.348	.430	.378
40	.373	.336	.324	.404	.359
30	.352	.312	.314	.392	.342
20	.314	.296	.300	.369	.320
10	.288	.268	.282	.334	.293

2. Cognition of Semantic Relations (CMR)

SG.	MAN	CDC	TO	O/A
.067	.072	.104	.038	.070
.052	.052	.070	.024	.050
.048	.047	.062	.017	.044
.036	.034	.048	.015	.033
.032	.031	.040	.012	.029
.026	.028	.034	.009	.024
.022	.018	.026	.006	.018
.022	.014	.022	.003	.015
.017	.009	.010	.001	.008

3. Memory for Semantic Units (MMU)

Percentile	SG	MAN	CDC	TO	O/A
90	.877	.866	.868	.884	.874
80	.866	.857	.852	.870	.861
70	.850	.846	.844	.858	.850
60	.842	.836	.836	.850	.841
50	.837	.822	.826	.844	.832
40	.825	.814	.819	.838	.824
30	.816	.809	.807	.821	.813
20	.800	.800	.798	.812	.802
10	.786	.772	.774	.800	783

4. Evaluation of Symbolic Implications (ESI)

SG	MAN	CDC	TO	O/A
1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	.991	.998
.992	.996	.996	.985	.992
.991	.992	.991	.982	.989
.981	.982	.982	.971	.979
.972	.960	.981	.955	.967
.943	.936	.972	.922	.943

*SG= Study Guides
MAN= Manuals
CDC= Career Development Course
TO= Technical Order
O/A= Overall

5. Convergent Production of Semantic Systems (NMS)

Percentile	*SG*	*MAN*	*CDC*	*TO*	*O/A*
90	.250	.250	.500	.500	.375
80	0.000	.250	.250	.250	.188
70	0.000	0.000	.250	.250	.125
60	0.000	0.000	0.000	.250	.062
50	0.000	0.000	0.000	.250	.062
40	0.000	0.000	0.000	0.000	0.000
30	0.000	0.000	0.000	0.000	0.000
20	0.000	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000

6. Convergent Production of Semantic Implications (NMI)

SG	*MAN*	*CDC*	*TO*	*O/A*
.579	.605	.580	.656	.605
.548	.550	.548	.581	.557
.525	.524	.523	.532	.526
.500	.520	.500	.524	.511
.500	.500	.478	.524	.500
.478	.478	.438	.490	.471
.444	.442	.415	.458	.440
.406	.420	.394	.438	.414
.382	.389	.368	.415	.388

7. Divergent Production of Semantic Units (DMU)

Percentile	*SG*	*MAN*	*DCD*	*TO*	*O/A*
90	.235	0.000	.464	0.000	.175
80	0.000	0.000	0.000	0.000	0.000
70	0.000	0.000	0.000	0.000	0.000
60	0.000	0.000	0.000	0.000	0.000
50	0.000	0.000	0.000	0.000	0.000
40	0.000	0.000	0.000	0.000	0.000
30	0.000	0.000	0.000	0.000	0.000
20	0.000	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000

8. Yngve Depth (YD)

SG	*MAN*	*CDC*	*TO*	*O/A*
.665	.649	.705	.684	.676
.645	.623	.649	.637	.638
.632	.610	.636	.622	.625
.624	.598	.613	.606	.610
.608	.581	.595	.578	.590
.594	.565	.571	.558	.572
.571	.552	.559	.552	.558
.548	.532	.534	.537	.538
.516	.506	.499	.512	.508

9. Morpheme Depth (MD)

Percentile	*SG*	*MAN*	*DCD*	*TO*	*O/A*
90	.759	.712	.732	.764	.742
80	.704	.666	.706	.706	.696
70	.668	.654	.690	.690	.676
60	.655	.644	.681	.671	.663
50	.640	.636	.652	.660	.647
40	.626	.612	.640	.643	.630
30	.612	.596	.628	.634	.618
20	.594	.584	.613	.617	.602
10	.560	.510	.600	.600	.568

10. Transformational Complexity (TC)

SG	*MAN*	*DCD*	*TO*	*O/A*
.994	.997	1.000	1.000	.998
.990	.989	1.000	1.000	.995
.986	.981	.994	1.000	.990
.979	.970	.992	.990	.983
.970	.966	.984	.988	.977
.967	.960	.980	.975	.970
.960	.950	.974	.971	.964
.947	.923	.960	.964	.948
.893	.873	.946	.957	.917

11. Center Embededness (CE)

Percentile	*SG*	*MAN*	*CDC*	*TO*	*O/A*
90	1.000	1.000	1.000	1.000	1.000
80	1.000	1.000	1.000	1.000	1.000
70	1.000	.833	.866	1.000	.925
60	.857	.800	.845	1.000	.876
50	.800	.690	.732	.800	.756
40	.667	.600	.667	.586	.630
30	.600	.500	.586	.464	.538
20	.500	.292	.333	.200	.331
10	.325	.000	.167	.000	.123

12. Left Branching (LB)

SG	*MAN*	*CDC*	*TO*	*O/A*
1.000	1.000	1.000	1.000	1.000
.944	1.000	1.000	1.000	.986
.833	.817	.775	1.000	.856
.800	.708	.667	.866	.760
.714	.500	.600	.690	.626
.589	.414	.500	.600	.526
.400	.387	.310	.464	.390
.333	.250	.167	.250	.250
0.000	0.000	0.000	0.000	0.000

13. Right Branching (RB)

Percentile	*SG*	*MAN*	*CDC*	*TO*	*O/A*
90	.400	.382	.412	.425	.405
80	.372	.333	.372	.392	.367
70	.358	.312	.357	.362	.347
60	.333	.297	.333	.342	.326
50	.324	.286	.308	.322	.310
40	.294	.250	.286	.312	.286
30	.274	.244	.256	.308	.270
20	.244	.233	.227	.286	.248
10	.224	.196	.208	.250	.220

14. Deleted Complement (DC)

SG	*MAN*	*CDC*	*TO*	*O/A*
1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000
.938	.800	.829	1.000	.892

☆ U.S. GOVERNMENT PRINTING OFFICE: 1977—771-057/18